

CROATIAN JOURNAL OF PHILOSOPHY

Political Philosophy

LARRY S. TEMKIN
JONATHAN SEGLOW
SEBASTIÁN RUDAS
MIKLOS ZALA
SERGIO FILIPPO MAGNI
MANUEL KNOLL
IVAN CEROVAC
NENAD MIŠČEVIĆ

Articles

RADIM BÉLOHRAD
PIERRE STEINER

Book Reviews

MARKO KONJOVIĆ
GREGORY FRISBY

Croatian Journal of Philosophy

1333-1108 (Print)

1847-6139 (Online)

Editor:

Nenad Miščević (University of Maribor)

Advisory Editor:

Dunja Jutronić (University of Maribor)

Managing Editor:

Tvrtko Jolić (Institute of Philosophy, Zagreb)

Editorial board:

Stipe Kutleša (Institute of Philosophy, Zagreb),

Davor Pećnjak (Institute of Philosophy, Zagreb)

Joško Žanić (University of Zadar)

Advisory Board:

Elvio Baccarini (University of Rijeka), Carla Bagnoli (University of Modena), Boran Berčić (University of Rijeka), István M. Bodnár

(Central European University), Vanda Božičević (Bergen Community College), Sergio Cremaschi (Milano), Michael Devitt

(The City University of New York), Peter Gärdenfors (Lund University), János Kis (Central European University), Friderik

Klampfer (University of Maribor), Željko Loparić (Sao Paolo),

Miomir Matulović (University of Rijeka), Snježana Prijic-Samaržija (University of Rijeka), Igor Primorac (Melbourne),

Howard Robinson (Central European University), Nenad Smokrović (University of Rijeka), Danilo Šuster (University

of Maribor)

Co-published by

“Kruzak d.o.o.”

Naserov trg 6, 10020 Zagreb, Croatia

fax: + 385 1 65 90 416, e-mail: kruzak@kruzak.hr

www.kruzak.hr

and

Institute of Philosophy

Ulica grada Vukovara 54/IV, 10000 Zagreb, Croatia

fax: + 385 1 61 50 338, e-mail: filozof@ifzg.hr

www.ifzg.hr

Available online at <http://www.ceeol.com> and www.pdcnet.org

CROATIAN
JOURNAL
OF PHILOSOPHY

Vol. XIX · No. 55 · 2019

Political Philosophy

Neutrality and the Relations between Different Possible Locations of the Good LARRY S. TEMKIN	1
Religious Accommodation: An Egalitarian Defence JONATHAN SEGLOW	15
Being a Progressive in Divinitia SEBASTIÁN RUDAS	37
Laborde's Liberalism's Religion: The Problem of Religious Exemptions MIKLOS ZALA	55
Procreative Beneficence toward Whom? SERGIO FILIPPO MAGNI	71
Michael Walzer's Republican Theory of Distributive Justice: "Complex Equality" as Equal Freedom from Domination MANUEL KNOLL	81
The Epistemic Value of Partisanship IVAN CEROVAC	99
Populists, Samaritans and Cosmopolitans. What is the Right Alliance? NENAD MIŠČEVIĆ	119

Articles

- On Three Attempts to Rebut
the Evans Argument against Indeterminate Identity
RADIM BĚLOHRAD 137
- Content, Mental Representation
and Intentionality: Challenging
the Revolutionary Character of Radical Enactivism
PIERRE STEINER 153

Book Reviews

- Clare Chambers, *Against Marriage:
An Egalitarian Defense of the Marriage-Free State*
MARKO KONJOVIĆ 175
- Tim Crane, *The Meaning of Belief:
Religion from an Atheist's Point of View*
GREGORY FRISBY 181

*Neutrality and the Relations between Different Possible Locations of the Good**

LARRY S. TEMKIN
Rutgers University, New Brunswick, USA

This article explores and challenges several common assumptions regarding what neutrality requires of us in assessing outcomes. In particular, I consider whether we should be neutral between different possible locations of the good: space, time, and people. I suggest that from a normative perspective we should treat space differently than time, and people differently than space and time. I also argue that in some cases we should give priority to people over space and time, and to time over space, but that, controversially, in some cases we should give priority to time over people.

Keywords: Neutrality, impartiality, practical reasoning, dominance principles, impersonal ideals, utilitarianism, space, time.

1. Introduction

It is common for philosophers and others to assume that, in certain contexts, morality requires us to be *neutral* with respect to space, time, and people. Following John Broome, we might refer to space, time, and people as different possible *locations* of the good (Broome 1991). Arguably, classical utilitarians would have insisted that all that mattered was the sum total of utility that obtained in the world, not its location. So, in principle, classical utilitarians would have insisted not only on neutrality *within* each of the different possible locational categories, but *between* each of these locational categories.

* This article was originally given as a talk at the Value Conference sponsored by The Ohio State University, Maribor University, and Rijeka University, held in Dubrovnik, Croatia, 11–15 June, 2018. I am grateful to the organizers, Justin D'Arms, Edin Lin, Boran Berčić, and Nenad Mišćević, for inviting me to participate in that conference. I am also grateful to Nenad Mišćević for inviting me to submit a written version of my talk for inclusion in this journal.

Setting aside the special relations that give rise to agent-relative prerogatives and duties, I think there is surely something to be said in favor of the general proposition that morality requires us to be neutral with respect to space, time, and people. However, I believe it is much less clear what that is than many have assumed. In particular, I believe that many common assumptions about the relative status of the different possible locations of the good are dubious.

In this article, I shall briefly try to illustrate some of my reasons for thinking this via a number of examples. The article is divided into three main sections. In Section 2, *Space and time*, I offer an example where I treat space and time differently for the purposes of rational deliberations. In Section 3, *Space, time, people, and dominance principles*, I introduce three dominance principles with respect to space, time, and people. I note that, intuitively, it may seem that if we should be neutral between the different possible locations of the good, then if we accept one of the dominance principles we should accept all three. However, I note that in cases involving infinity the three dominance principles conflict, so that we *cannot* accept all three. This suggests that we must reject all three dominance principles, restrict their scope, or reject neutrality between the three different locations of the good. In Section 4, *On the independent normative significance of temporal value*, I suggest that there may be independent value to filling different temporal periods with high quality sentient life, beyond the extent to which doing so is good or bad *for* the sentient beings in those time periods.

Together, my arguments suggest that, from a normative perspective, we should treat space differently than time, and people differently than space and time. More specifically, they suggest that in some cases we should give priority to people over space and time, and to time over space, but that, controversially, in *some* cases, we should, in essence, give priority to time over people.

2. *Space and time*

Let's start with the question of whether, from a normative perspective, we must treat space and time the same. Some people who do philosophy of physics may think that to even raise this question reveals a deep misunderstanding of the nature of space and time, since modern physics supposedly tells us that space and time are inextricably linked in a single space/time continuum. However, my concern, here, is with whether we *must* treat space and time the same, in virtue of some *necessary normative* principle, not with the *empirical* relation between space and time that may (or may not) *in fact* obtain in our world, but need not obtain in all possible worlds. Accordingly, to help us think about the questions that concern me here, we should assume, throughout this article, that we live in a world where a non-relativistic conception of space and time is true. This would have been the case if either a Newtonian or a pre-Newtonian conception of space and time had turned out to be true.

Bearing the preceding assumption in mind, consider the following thought experiment.

Suppose I learn that our civilization will live in our galaxy another 1000 years, and then die out. I *also* learn that in a distant galaxy another civilization will exist for the same 1000 years and then die out. Suppose, further, that this is also so in a third distant galaxy, and a fourth. I find this all quite interesting. Perhaps oddly, it is somewhat *pleasing* to me to learn that there are, in fact, advanced civilizations living in various galaxies far away. However, suppose I also learn that *beyond* the fourth galaxy there is *nothing* but cold, empty, space. This, too, I find interesting. Yet, I must confess that learning that fact doesn't bother me very much at all. Indeed, if someone said that events beyond the fourth galaxy were about to unfold which would make those distant reaches inhospitable to life forms in perpetuity, I wouldn't think it especially important for our civilization to make significant sacrifices, if it could, to prevent that from happening.

Suppose, on the other hand, I vary the story a bit. As before, I learn that civilization in our galaxy will die out in 1000 years, but I learn that after ours dies out another, wholly unrelated, civilization will arise and persist for 1000 years in a second galaxy far away, and that this will happen again a third and fourth time. But after that, I am told, there will be *nothing* but cold, empty, space, *forever*. For some reason, *that* knowledge would bother me a *lot*. Indeed, if I learned that events were about to unfold which would make the universe uninhabitable for any life forms 4000 years from now, unless our civilization made significant sacrifices to prevent that from happening, I would feel quite strongly that we should do so. Moreover, I would feel that way even if I knew that *our* civilization was going to die out in 1000 years no matter what we did, and that any future civilizations would do nothing to advance or realize our particular dreams or goals.¹

¹ Samuel Scheffler (2013) has argued that having descendants who will help realize some of our deepest hopes, projects, or ideals, helps to give our lives value and meaning that they would otherwise lack. Scheffler's views are entirely compatible with my own, and I am happy to accept them. But, they point to *other* reasons why one might be more concerned about the future than about what happens elsewhere in space than those I am trying to illuminate here. As my example makes plain, I believe that even if the future civilizations were wholly unrelated to our own, and would do *nothing* to further *our* particular hopes, projects, and ideals, I *still* believe that there would be strong reason to ensure that such civilizations would exist if they would have high quality lives. In addition, I believe that such reasons would be stronger than any we would have to ensure, were it possible, that such civilizations obtain elsewhere in space contemporaneous with our own.

Similarly, Jeff McMahan (personal communication, October 2, 2015) suggested a variety of considerations that might lead us, in general, to give greater weight to there being high quality sentient lives existing in the future, than to there being high quality sentient lives existing elsewhere in space. According to McMahan, these might include views we have about the importance of the preservation of value, views about the importance of progress, and views about the importance of greater diversity of experiences. My response to McMahan is threefold.

My views here may ultimately be indefensible, but I don't think they are idiosyncratic, and they reveal that I have an asymmetry in my thinking about space and time. Even if I know that the exact same number of sentient beings will exist in the two scenarios, and will flourish to the exact same extent, I respond to the two scenarios very differently. The second scenario seems very bad indeed. The first seems hardly bad at all. Thus, I think it very important that many *times* be filled with flourishing sentient beings. I think it much less important that many *spaces* be filled with flourishing beings. Of course, these reflections hardly constitute an *argument* for my view; but they reveal that I am treating space and time differently in my moral deliberations.

Let us vary the preceding example just a bit. Suppose that we had taken steps to ensure both that three distant planets in space were populated with advanced civilizations, and that each of the next three consecutive hundred year periods would also be populated with advanced civilizations. Suppose, next, that someone developed two pills, each of which would enable us to lead really flourishing lives for 120 years, but with the following consequences. If we take the red pill, we won't be able to populate any more distant planets in space. If we take the blue pill we won't be able to populate any more hundred year periods after the next three. I believe that there would be little or no objection to taking the red pill, but *very* strong objection to taking the blue one. Here, again, I find myself wanting to treat space and time differently in my moral deliberations.

As discussed in note three, the objection to taking the blue pill that I have in mind here is independent of any of the ways in which future civilizations might help realize *our* projects or goals. To my mind, a *principle* reason for wanting high quality life to exist in the future is also a reason for wanting high quality life to exist in the past; namely that it is a very good thing for different time periods to be filled with a significant number of sentient beings with high quality lives. *That* reason is unrelated to any meaning or value that the existence of descendants may sometimes help bestow on their ancestors. I don't believe that there is a similarly strong reason to fill different locations in space with high quality lives.

First, as with what I said about Scheffler's view, I don't regard my position as incompatible with McMahan's. Depending on the details of the case, there could be more than one reason for valuing the existence of *future* civilizations over the existence of contemporaneous civilizations elsewhere in space. But second, in my examples, I wasn't, in fact, assuming that there was greater diversity of experiences over time than across space, nor was I assuming that there would be *progress* between our current civilization and the future, unrelated, civilizations. Thus, my views about such cases weren't, in fact, turning on such factors. Moreover, importantly, I note that the notions of *preservation* of value, and *progress*, have a temporal dimension built in to them, but not a spatial dimension. So, McMahan's suggestions regarding those factors would, if correct, not be a *rival* to my own, but rather a further elucidation of some of the *reasons* why we should treat space and time differently for the purposes of practical reasoning.

3. *Space, time, people, and dominance principles*

Suppose, for a moment, that we should dismiss my worries, and go along with the widely-held view that, except in the case of special relations and special obligations, we should be neutral across all three dimensions of space, time, and people. A corollary of such a view would *seem* to be that if we accept a dominance principle with respect to one of these categories, we should accept a similar dominance principle with respect to the others. Consider, for example, the following three dominance principles regarding utility.

Spatial Dominance Principle: for any two alternative outcomes, A and B, if A and B involve the same region of space, S, which is made up of a set of non-empty sub-regions of space, s_1, \dots, s_n , if A is better than B regarding utility in *every* sub-region of space, s_i , then A is better than B regarding utility.

Temporal Dominance Principle: for any two alternative outcomes, A and B, if A and B involve the same region of time, T, which is made up of a set of non-empty sub-regions of time, t_1, \dots, t_n , if A is better than B regarding utility in *every* sub-region of time, t_i , then A is better than B regarding utility.

Personal Dominance Principle: for any two alternative outcomes, A and B, if A and B involve the same people and A is better than B regarding utility for *every* person who will ever live, then A is better than B regarding utility.

Intuitively, many would find each of the preceding dominance principles plausible, and they might assume that if one of them is true the others must also be true. But this assumption is clearly false. To see this, consider Diagram 1.²

² The following case is my own, but it was sparked by an example I first heard during a discussion with John Broome, many years ago, which he called “Expanding Heaven and Expanding Hell.” Broome credited his example to James Cain (See Cain 1995). Although my views about this topic were arrived at independently, other philosophers have developed similar arguments in order to make similar points. See, for example, Vallentyne (1993), Lauwers (1997) Vallentyne and Kagan (1997), Machina (2000), Lauwers and Vallentyne (2004), Bostrom (2011), and Campbell (2015).

Interestingly, while Cain uses an example similar to mine to arrive at the same conclusion that I do regarding the relative status of Personal and Temporal Dominance Principles for certain cases and contexts, Campbell produces a series of ingenious examples in order to show that, depending on one’s theory of personal identity, there will be other cases where the relative status of Personal and Temporal Dominance Principles would be the reverse of that for which Cain and I argue. I don’t favor the reductionist view of personal identity that would lead to Campbell’s results. However, many do, and for those who do, Campbell’s arguments are quite compelling.

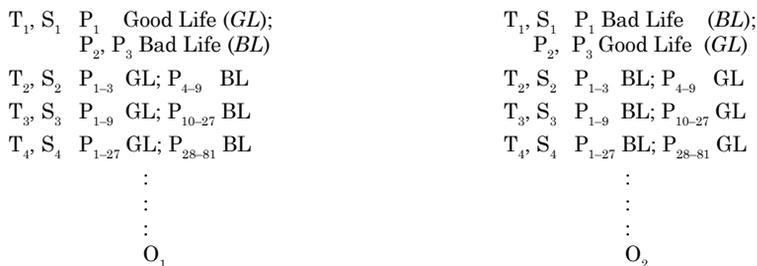


Diagram 1.

O_1 and O_2 are two possible outcomes, in which the very same people, P_1, P_2, P_3 , etc. exist. In O_1 there is one person, P_1 , living in time period one, and spatial region one, who has a *good* life, well *above* the level at which life ceases to be worth living; but, unfortunately, there are *twice* as many *other* people, P_2 and P_3 , who have *bad* lives, well *below* the level at which life ceases to be worth living. In time period two, P_1 through P_3 have moved to spatial region two, where they all enjoy good lives; but, unfortunately, in that time period, and at that location, twice as many other people, P_4 through P_9 , have come into existence, and their lives are as bad as P_2 and P_3 's lives were during T_1 . In time period three, P_1 through P_9 have all moved to spatial region three, where they all enjoy good lives, but, unfortunately in that time and location, twice as many other people, P_{10} through P_{27} , have come into existence, and their lives are as bad as P_2 and P_3 's lives were during T_1 . Outcome One continues to unfold, in this ever expanding manner, forever, with each time period lasting for one day, and each person living for 100 years total, before dying. Here, and later, we assume that the positive value of each good moment is the same, the negative value of each bad moment is the same, and that the two values sum to zero. So, by hypothesis, a life containing an equal number of moments of good and bad life will have a net value of zero, a life containing more moments of good life than bad will have a positive net value, and a life containing more moments of bad life than good will have a negative net value.

Outcome Two is analogous to, though the reverse of, Outcome One. In O_2 , there is one person, P_1 , living in time period one, and spatial region one, who has a *bad* life, well *below* the level at which life ceases to be worth living; but, fortunately, there are *twice* as many *other* people, P_2 and P_3 , who have *good* lives, well *above* the level at which life ceases to be worth living. In time period two, P_1 through P_3 have moved to spatial region two, where they all suffer bad lives; but, fortunately, in that time period, and at that location, twice as many other people, P_4 through P_9 , have come into existence, and their lives are as good as P_2 and P_3 's lives were during T_1 . And so on. As before, Outcome Two continues to unfold, in this ever expanding manner, forever, with each time period lasting for one day, and each person living for 100 years total, before dying.

How do Outcomes One and Two compare regarding utility? Comparing them spatial location by spatial location, or temporal location by temporal location, Outcome Two would be clearly *better* than Outcome One, in accordance with the Spatial and Temporal Dominance Principles. This is because for *every* spatial region, S_n , and *every* temporal region, T_n , there will be twice as many people with good lives as with bad lives in Outcome Two, while there will be twice as many people with bad lives as with good lives in Outcome One.

So, *should* we conclude that Outcome Two really *is* better than Outcome One regarding utility? I think not. This is because Outcome One is *better* than Outcome Two in accordance with the *Personal* Dominance Principle regarding utility. After all, by hypothesis, the *same* people exist in both outcomes, and they are all *clearly* better off in Outcome One, where they each suffer for only *one* bad *day* followed by *99 years* and *364* days of good life, than they are in Outcome Two, where they each fare well for only *one day*, followed by *99 years* and *364* days of bad life.

In this example, we can accept the dominance principle regarding people, or we can accept the dominance principles regarding space and time, but we cannot do both. Here we have a proof that, unless we reject *all three* dominance principles, in some cases, at least, we *should* not, and *cannot*, treat space and time the same way as we treat people. So, *should* we reject all three dominance principles? I don't see why. In this case, at least, the Personal Dominance Principle clearly seems to support the correct answer!

The preceding argument suggests that for certain cases, at least, we should give priority to distributions of wellbeing across people over distributions of wellbeing across space and time. And earlier, I suggested being more concerned about distributions of wellbeing throughout time, than throughout space. The priority ranking of people, over time, over space, for some cases, at least, might be further buttressed by considering Diagram 2.

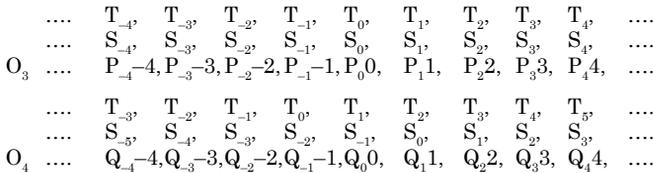


Diagram 2.

In Diagram 2, O_3 and O_4 are outcomes with an infinite number of people, P_i or Q_j , with each person, located at a particular location in space, S_k , and a particular location in time, T_l , at a level corresponding to one of the integers. So, for example, in Outcome Three, person P_0 is at level 0, at temporal location 0 and spatial location 0, while in Outcome Four person Q_{-4} is at level -4 , at temporal location -3 and spatial location -5 . For the purposes of this example, I am assuming that the metaphys-

ics of space and time allow for the identification of the same spatial and temporal locations across different possible outcomes, so that for each k and l , S_k corresponds to the very same location in space in each outcome, and T_l corresponds to the very same location in time in each outcome. If such an assumption is coherent, then Diagram 2 illustrates that the Spatial and Temporal Dominance Principles are incompatible with each other. So we can reject both, or limit their scopes, but we cannot simply accept both. This is because, as a careful examination reveals, in Diagram 2, Outcome Three is *better* than Outcome Four at every point in *time*, but Outcome Three is *worse* than Outcome Four at every point in *space*.

Now assume, temporarily, that the populations of the Outcomes Three and Four are wholly distinct. In that case, I can see why someone might claim that each outcome is *equally* good, since each involves an infinite number of people, such that for each integer there is exactly one person whose level of wellbeing is accurately represented by that integer. If one reasoned in that way, then one would be rejecting *both* the Spatial Dominance Principle *and* the Temporal Dominance Principle. However, my own judgment is that in this case we should *accept* the judgment yielded by the *Temporal* Dominance Principle, and *reject* the judgment yielded by the *Spatial* Dominance Principle. That is, in this case, I would judge Outcome Three as better than Outcome Four, since it is better at each moment in time and, to my mind, there is neither a compelling reason to ignore this consideration, nor a countervailing reason outweighing it.

However, as the previous discussion makes plain, I believe that there *could* be a countervailing reason outweighing, or perhaps undermining, the Temporal Dominance Principle. In particular, if the *same* people would exist in both Outcome Three and Outcome Four, and they would *each* be better off in one of the outcomes than the other, then, in accordance with the *Personal* Dominance Principle, I would judge, in *this particular* case and context, the outcome in which they were all better off as better than the other outcome, regarding utility, regardless of how the two outcomes compared in accordance with either the Spatial or Temporal Dominance Principles.³

³ As the literature cited in note four reveals, many people have recognized that Dominance Principles fail in infinite cases. And many others are suspicious of appealing to infinite cases in thinking about normative issues. Given the difficulty of grasping the infinite intuitively, the latter attitude is understandable. Nevertheless, I think it is deeply mistaken. I believe that if one is careful, one can usefully consider infinite cases when doing normative philosophy, and that there can be great philosophical payoff from doing so. I also believe that since it is very possible that we live in an infinite universe, it would be deeply problematic if our moral principles were only plausible for, and applicable to, finite realms.

Unfortunately, the issues connected with this topic are too complex to pursue here. Still, I believe that the infinite examples canvassed in this work are appropriate for the purposes to which I put them, and that we can usefully gain insight into this article's topics by considering them. I might add that many people assume

4. *On the independent normative significance of temporal value*

Many philosophers, economists, and others believe that the proper locations of value are *people*, and that considerations of time and space are *only* relevant insofar as they have an effect on the quality of different people's lives. But the considerations I have offered in support of giving greater priority to time over space belie that simple, natural, assumption. Elsewhere, I have argued at great length in support of there being *impersonal* ideals, as well as *personal* ones.⁴ In particular, I have argued that many of the ideals that people value most, including such ideals as justice, equality, beauty, perfection, and truth, have value *beyond* the extent to which they are good or bad *for* people. I have also argued in favor of a *Capped Model of Ideals*, according to which there may be an upper limit to how good an outcome can be, regarding utility, for any given period of time, so that once a large number of people exist who are already extremely well off, merely adding more people to the outcome with lives worth living won't significantly make the outcome better.⁵ I can't repeat the arguments for these positions here. However, if, in fact, they are correct, they provide the theoretical basis for a rather surprising, and controversial, conclusion. To wit: while in the contexts previously discussed there was reason to give priority to *people over times*, in *some* contexts, there may be reason, in essence, to give priority to *times over people*.

To see this, consider Diagram 3.

that even if the Dominance Principles fail in infinite cases, they succeed in finite cases. However, I believe that this intuitively plausible position is also mistaken, for reasons that I have given elsewhere and won't repeat here. For reasons relevant to rejecting the Personal Dominance Principle, even in finite cases, see Temkin (1993, 2000: 126–161, 2003a; 2003b, and 2012). For reasons relevant to rejecting the Spatial and Temporal Dominance Principles, see Temkin (2012 and 2015).

⁴ See my *Inequality*, "Equality, Priority, and the Levelling Down Objection," "Egalitarianism Defended," "Personal versus Impersonal Principles: Reconsidering the Slogan," and *Rethinking the Good*.

⁵ At least if the additional people are no better off than those who already exist. I was initially led to advocate a Capped Model of Ideals by reflecting on Derek Parfit's *Repugnant Conclusion* (Parfit 1984: 388). For a detailed explication of the Capped Model of Ideals, and some of the considerations underlying it, see Chapter 10 of *Rethinking the Good*.

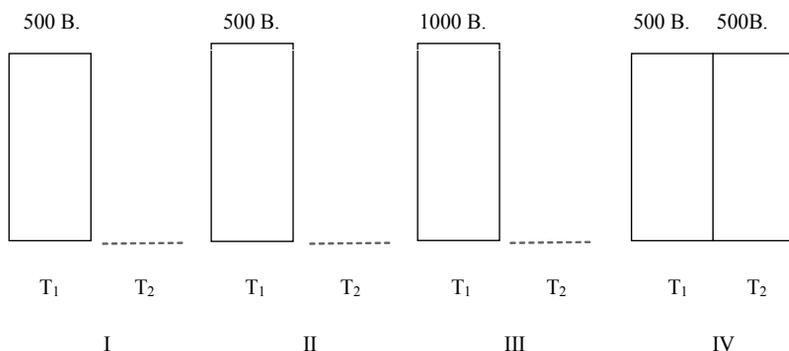


Diagram 3.

Diagram 3 is different from Diagrams 1 and 2 in one important respect. In Diagram 3, the width of each rectangle represents the length of a *time* period, rather than the number of people existing in an outcome—that number is given above each rectangle. Moreover, we assume, for the sake of the example, that in accordance with a Capped Model for Ideals, for any time period the duration of T_1 , 500 billion people, at any given level of wellbeing, would be enough to bring an outcome *very near* to the cap for how good an outcome can be, during that time period, by having lots of people with that level of wellbeing.

As drawn, Diagram 3 represents four possible outcomes. In Outcome I, there are 500 billion people all at a very high level, spread out over a lengthy time period, T_1 . Unfortunately, in Outcome I, a second equally lengthy time period, T_2 , is *utterly devoid* of high quality sentient life. In Outcome II, the very same 500 billion people who would exist in Outcome I exist at a *slightly* higher quality of life during time period T_1 . However, once again, unfortunately, time period T_2 is *utterly devoid* of high quality sentient life. In Outcome III, an extra 500 billion people have been added to the temporal period T_1 , at the same level as those in Outcome II. But, once again, unfortunately, time period T_2 is *utterly devoid* of high quality sentient life. Finally, in Outcome IV, the original 500 billion people occupy time period one at the level of those in Outcome I, and there are an extra *500 billion people* who all exist *in time period* T_2 at the very same level. Moreover, by hypothesis, these are the very same “extra” people who *would* have existed in Outcome III during time period T_1 at a slightly higher level.

How do the different outcomes compare, all things considered? Drawing on the results presented previously, together with my view about how best to understand and interpret the Capped Model of Ideals—which I have only had time to barely *mention*, but not develop in this article—I would make the following judgments. I would judge Outcome II as better than Outcome I, since it is better for everybody. However, since it is only a *little* bit better for everyone, and I reject

a simple additive aggregationist model for ranking outcomes, I would judge that Outcome II isn't a *whole* lot better than Outcome I.

Next, I would judge Outcome III as only slightly better than Outcome II, and hence, like Outcome II, as only a little better than Outcome I. This is an implication of the Capped Model of Ideals, given our assumption that the cap is already almost reached in Outcome II, for how good an outcome can be by having lots of people at that level, during a time period the duration of T_1 . On that assumption, the Capped Model implies that merely adding another 500 billion people to the *very same temporal region* would *not* make a significant difference to the overall goodness of an outcome. Hence, as indicated, Outcome III would not be significantly better than Outcome II.

On the other hand, since I believe that it *is* very important that many regions of *time* be filled with flourishing beings, I would judge Outcome IV, where there are *500 billion flourishing beings in T_2* as *much* better than Outcome I, where T_2 is *utterly devoid* of high quality life. Unsurprisingly, then, I would also judge Outcome IV as better than Outcome's II and III, which are only a little better than Outcome I in terms of what happens during time period T_1 , and which, like Outcome I, have the significant negative feature of being utterly devoid of high quality sentient life throughout the whole of time period T_2 .

But notice, by hypothesis, the *very same people* exist in Outcomes III and IV, and they are *all* better off, even if only slightly, in Outcome III than in Outcome IV. So, my judgment that Outcome IV is better than Outcome III, all things considered, suggests that, in *this* context, I am, as it were, giving priority to time over people. This is a striking conclusion that many people will initially find deeply implausible.⁶ However, I believe that this conclusion is defensible, and that, on reflection, it is neither surprising nor implausible. It is merely yet another manifestation of the fact that some ideals have *impersonal* value, in the sense that their realization can contribute to the goodness of outcomes, *beyond* the extent to which their realization is good *for* the sentient beings in those outcomes. In particular, this article's considerations reflect the view that there can be significant impersonal value to filling different periods of time with high quality life. Correspondingly, it shouldn't be surprising that in some cases, such as the one depicted by Diagram 3, such impersonal value can outweigh the personal value of increasing individual wellbeing by a small amount.

⁶ Derek Parfit once referred to a claim of this sort as *The Absurd Conclusion* (Parfit 1984: 410–411). Arguably, underlying the plausibility of Parfit's ascription were both a *welfarist* view—which assesses the goodness of outcomes solely in terms wellbeing, thereby rejecting the relevance of *impersonal* ideals for assessing outcome goodness—and a standard view about what neutrality requires. As my claims here make plain, I believe that both views are dubious. For further arguments in support of this article's claims and the view that The Absurd Conclusion isn't, in fact, absurd, see Temkin (Forthcoming).

5. Conclusion

Let me briefly summarize my main claims. Setting aside any agent-relative duties and permissions that may arise due to the special relations that sometimes obtain among different beings, most people agree that morality requires neutrality. There is, I believe, something deeply right about this. However, it is much less clear than many have supposed what neutrality entails.

In particular, one might have presumed that morality requires us to be neutral both within, and between, each of the different possible locations of good: space, time, and people. While there may be *some* sense in which this is true, I have suggested that we need not, and should not, treat space, time, and people the same for the purposes of normative reasoning.

I have offered examples suggesting that in some cases we should give priority to time over space, and that it is more important that different temporal regions be filled with flourishing sentient beings than that different spatial regions be filled with flourishing beings. I have also shown that three intuitively plausible dominance principles conflict in certain cases: a Spatial Dominance Principle, a Temporal Dominance Principle, and a Personal Dominance Principle. I have suggested that when they conflict I favor the judgment of the Temporal Dominance Principle over that of the Spatial Dominance Principle, and that in certain cases I favor the judgment of the Personal Dominance Principle over that of both the Spatial and Temporal Dominance Principles. However, drawing on claims argued for elsewhere—that there can be impersonal ideals, as well as personal ideals, relevant for assessing outcome goodness, and that we need something like a Capped Model of Ideals for evaluating outcomes—I also argued for the striking claim that in certain cases we should, as it were, give priority to times over people.

This article is very much a preliminary exploration, and I am acutely aware that the speculative lines I have pursued will strike many as wild, implausible, and deeply wrongheaded. Nevertheless, I believe there is much to be learned about the nature of practical reasoning by taking these issues seriously; even if doing so may ultimately take us in directions other than those that I have stumbled towards in this article.⁷

References

- Bostrom, N. 2011. "Infinite Ethics." *Analysis and Metaphysics* 10: 9–59.
 Broome, J. 1991. *Weighing Goods*. Oxford: Basil Blackwell.

⁷ Most of this article's issues have been dealt with in much greater depth in Temkin (2012, 2015 and forthcoming). Even so, our thinking about these topics remains very much in an early stage, and there is much more to be learned, and written, about all of these issues.

- Cain, J. 1995. "Infinite Utility." *Australasian Journal of Philosophy* 73: 401–4.
- Campbell, T. 2015. *Personal Ontology and Bioethics*. PhD diss. Rutgers University.
- Lauwers, L. 1997. "Infinite Utility: Insisting on Strong Monotonicity." *Australasian Journal of Philosophy* 75 (2): 222–233.
- Lauwers, L. and Vallentyne, P. 2004. "Infinite Utilitarianism: More is Always Better." *Economics and Philosophy* 20: 307–330.
- Machina, M. 2000. "Barrett and Arntzenius's Infinite Decision Puzzle." *Theory and Decision* 49: 291–295.
- Parfit, D. 1984. *Reasons and Persons*, Oxford: Clarendon Press.
- Scheffler, S. 2013. *Death and the Afterlife*. New York: Oxford University Press.
- Temkin, L. Forthcoming. *Population Ethics Forty Years On: Lessons to be Learned from "Box Ethics"* (unpublished, 2017).
- Temkin, L. 1993. *Inequality*. New York: Oxford University Press.
- Temkin, L. 2000. "Equality, Priority, and the Levelling Down Objection." In M. Clayton and A. Williams (eds.). *The Ideal of Equality*. New York: Macmillan and St. Martin's Press.
- Temkin, L. 2003a. "Egalitarianism Defended." *Ethics* 113: 764–782.
- Temkin, L. 2003b. "Personal versus Impersonal Principles: Reconsidering the Slogan." *Theoria* 69: 20–30.
- Temkin, L. 2012. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. New York: Oxford University Press.
- Temkin, L. 2015. "Rationality with Respect to People, Places, and Times." *Canadian Journal of Philosophy* 45: 576–606.
- Vallentyne, P. 1993. "Utilitarianism and Infinite Utility." *Australasian Journal of Philosophy* 71: 212–217.
- Vallentyne, P. and Kagan, S. 1997. "Infinite Value and Finitely Additive Value Theory." *Journal of Philosophy* 94: 5–26.

Religious Accommodation: An Egalitarian Defence

JONATHAN SEGLOW
Royal Holloway, University of London, UK

This paper offers a distinctively egalitarian defence of religious accommodation in contrast to the rights-based approaches of contemporary legal thinking. It argues that we can employ the Rawlsian idea of a fair framework of co-operation to model the way that accommodation claimants reason with others (such as their employers) when they wish to be released from generally applicable rules. While participants in social institutions have 'framework obligations' to adhere to the rules those institutions involve, they also have 'democratic obligations' to re-consider and on occasion revise those rules which set back participants basic interest, including individuals' interest in manifesting their religion or belief. A number of objections to accommodation are considered, and it's argued that the personal responsibility objection is most serious. It's argued that responsibility can be interpreted through the notion of identification which in turn can be conceptualised through the ideal of integrity, and that the value of integrity in fact counts in favour of accommodation claims. The paper also offers replies to other objections to religious accommodation including the problem of proliferation, the problem of illiberal beliefs and the rewarding the doctrinaire objection.

Keywords: Religious accommodation, fairness, framework of co-operation, Rawls, personal responsibility, integrity.

1. *Introduction*

Across Europe and North America, the problem of religious accommodation continues to be the subject of legal debate, political campaigns and seemingly intractable philosophical argument. Whether, and if so, when with ordinarily applicable laws and rules should accommodate citizens with strong religious or moral convictions raises challenging questions about the meaning, interpretation and justification of those laws and rules. The problem is often conceptualised in terms of rights

and liberties. In a US context, this stems from the free exercise clause of the First Amendment. In a European context, the legal debate around accommodation revolves around article 9 of the European Convention of Human Rights (ECHR) which protects freedom of religion or belief. Both these documents conceptualise religious liberty as a right of pre-eminent weight and authority. Over the years an enormous case law (and associated commentary by academic lawyers) has developed over the proper interpretation, application and enforcement of both Article 9 and the First Amendment.

In what follows, however, I shall depart from this liberty- and rights-based perspective to defend a distinctively egalitarian defence of religious accommodation in a stronger sense of ‘egalitarian’ than has been employed in the literature so far.¹ My argument for accommodation not only assumes that both appropriate non-religious as well as religious convictions should be accommodated, or that we should distribute the burdens of accommodation in a fair and equitable way. It departs from the judicial model of accommodation where courts and tribunals insert their judgments into disputes between claimants, to conceptualise those disputes horizontally as ones between contending parties who deliberate over the rules that govern their interactions. In making this argument, I borrow from Rawls the ideal of society as a fair system of co-operation between free and equal citizens, though I suggest that that basic idea applies also to particular institution within society. I also give serious consideration to the objections to accommodation, and I suggest that the most powerful objection to accommodation arises from the notion of personal responsibility; roughly put, the idea that if we are responsible for our deeply held religious and moral convictions then we cannot in justice claim any special accommodation for them. The other main claim I make in this paper is that reflection on the ideal of personal responsibility in fact motivates an argument that counts in favour of, and not against, accommodation.

The plan of the remainder of the paper is as follows. In the rest of this Section 1 make a couple of preliminary comments. Section 2 sets out the egalitarian argument for a framework of social co-operation in general terms. I distinguish there between individuals’ duties to pursue their religious and other convictions within a fair framework of co-operation and their democratic duties to reflect with their fellow individuals on the fairness of that framework. Section 3 explores the

¹ For example Cécile Laborde describes ‘egalitarian theories of religious freedom’ as those which do not see religion as uniquely special; that do not protect religious commitments qua their religiosity; and that accord equal civic status to all citizens. See Laborde (2014: 53–8) and (2017: Part I). See also Shorten (2010) for an egalitarian defence of exemptions based on an ideal of equal citizenship. My own defence of accommodation employs a notion of egalitarianism stronger than both Laborde’s and Shorten’s insofar as it emphasises an ideal of participants in social co-operation who enjoy equal status. For a critique of the idea that accommodation has much to do with equality, see Jones (2017).

objections to accommodation. I distinguish there between internal objections to the very idea of accommodation and the external costs which any particular accommodation imposes, and outline four internal objections of which the personal responsibility objection is one. Section 4 argues in response that personal responsibility can be conceptualised in terms of identification, while Section 5 argues that identification is captured by an ideal of integrity. Since integrity has impersonal moral value, it in fact supports accommodation. Section 6 employs the notions of integrity and a fair framework of co-operation to offer some replies to the other internal objections. Section 7 returns to the ideal of social co-operation to show how individuals' interests in personal integrity as well as other interests can be put to work in a practical framework for examining accommodation claims which also takes account of external costs, while the final Section briefly concludes.

The topic of this paper is accommodation, not legal exemptions. Though the two are sometimes conflated, the latter is a broader idea. For example, if a cafeteria offers kosher and halal meat then it accommodates the religious preferences of its Jewish and Muslim customers. Such an accommodation might well draw on a legal exemption. In the UK, for example, Jews and Muslims enjoy a specific exemption from the laws on animal welfare which normally require that animals such as cows and lambs are stunned before they are slaughtered.² But most cases of accommodation do not involve a formal legal exemption. This is because most (though not all) cases of accommodation occur in employment, which is not a coincidence since the world of work imposes on employees rules and regulations which come into conflict with their religious, and less commonly non-religious, convictions. This fact has in my view been under-appreciated, by political theorists, though perhaps not by academic lawyers (see for example Vickers 2016). This may be due to a relative neglect by liberal political theorists of the sphere of work, possibly due to the influence of the Rawlsian idea that individuals pursue their abstractly defined conception of the good life.

A second preliminary comment is that my interest here is in individual, not group-based accommodation claims. Of course, an accommodation may be enjoyed by a number of individuals who are all part of a group (such as Christian Sabbatarians for example), and so too an exemption, as the example of the animal slaughter exemption in the UK makes clear. But both those kinds of accommodation are distinct from when a group qua group exercises a liberty, as Bob Jones University did when it excluded would be black students or as the firm Hobby Lobby did when it sought an exemption from President Obama's Affordable Care Act.³ Many of the moral issues in individual and group-based accommodation are the same; in particular, in both cases we need some

² The Welfare of Animals (Slaughter or Killing) Regulations 1995.

³ *Bob Jones University v. United States* 461 U.S. 574 [1983]; *Burwell v. Hobby Lobby* 573 U. S. [2014].

compelling reason why one agent should enjoy a privilege not extended to others: this is the central puzzle of accommodation. But group cases involve further issues to do with the identity of the group which there is not the space to consider here, albeit what I say has some relevance to them.⁴

2. *Fair terms of co-operation*

Imagine a Christian employee approaching his manager and requesting a release from the Sunday work rota in order that he can attend church. He will need to present reasons to his manager for this special treatment, and if the manager turns down his request she will need to explain to the employee why his claim cannot be accepted. The two parties discuss the matter together interpersonally, and each offers reasons which s/he hopes the other will accept. If the employee's reasons prevail in their deliberations, then the work rota will be revised, with probable knock on affects for other employees who may need to take up some of the slack. Thus in deliberating with her employee the manager is (or should be) representing their interests too. By contrast, if the employee accepts the manager's arguments that release from Sunday working is an unjustifiable special privilege, then he will either need to knuckle down and work on Sunday, even though his conscience tells him he should be at church, or else resign his job.⁵

The framework for evaluating accommodation claims that I propose in this paper models this simple two person deliberation. It captures the sense in which parties to accommodation claims approach each other horizontally, in contrast to the vertical way in which court and tribunal judgements are inserted into a dispute between contending parties. And just as it is better if accommodation disputes can be resolved amicably by the parties concerned, so I think it is also better normatively to conceptualise the puzzle of accommodation in this horizontal, interpersonal way. Although accommodation disputes do involve parties who are superordinate over others—managers and employees, for instance, or teachers and schoolchildren—there is a fundamental sense in which contending parties are equals who must reason together. They do so in what, borrowing from Rawls, I shall call a framework of social co-operation (Rawls 1993: 15–22). This is the set of laws and rules which give individuals a particular pattern of freedoms, opportunities, duties and prohibitions, and which they have a general responsibility to maintain. The laws of a liberal polity are one example of a framework of social co-operation, but so too are the rules of a particular organisation such as a firm insofar as they require individuals to meet the role-related duties their job entails. For now, I shall describe the conception of a

⁴ For an excellent analysis of group accommodation, see Shorten (2015). For a critique of the Supreme Court's decision in *Hobby Lobby*, see Cohen (2015).

⁵ A third option is to revise his belief in the necessity of attending church.

framework of co-operation in abstract and ideal terms; I shall add more concrete detail to the conception later, in Section 7.

The notion of a framework of social co-operation regards individuals as both co-authors of and subject to an institutional structure which regulates their interactions. A *fair* framework of co-operation enables individuals to meet their aims and realise their interests better than any alternative framework (or no framework at all). The framework of co-operation is in an idealised sense, the common possession of the individuals to whom it applies. Not a charter which they inherit from on high, it expresses their equal status as members of an institution. If unfair, it is likely to breed feelings of resentment and alienation among the losing parties. If fair, it gives individuals a motivation to meet the duties it entails. I emphasise the idea that the framework of social co-operation is a regulatory ideal which is interpersonal and which structures and shapes the interactions of individuals in their particular roles as employees, students, club members and of course, as citizens in the polity. Individuals have a general obligation to do their fair share in keeping the framework of social co-operation in place, meeting the duties and accepting the limits it imposes. This includes the duty to adjust their aims so that they are realisable within the framework as it stands. Call the set of duties each individual has to help maintain the framework of social co-operation, her framework obligations. This is the aspect of the framework of social co-operation which is emphasised by Rawls.

However, the accumulation of laws and rules in actual societies in practice reflects the customs, traditions and historical pedigree of those societies in a way which often unfairly serves its majority's (or at least some citizens') interests at the expense of others. Moreover, particular frameworks are arrived at by flesh and blood human beings, with normal biases and prejudices and in contingent circumstances which vary from place to place. As a result, particular frameworks of social co-operation may be unfair, in various ways. Where this is so, the individuals who are burdened by that unfairness benefit do not have a framework obligation to meet all the duties the framework imposes, at least not those which impose unfair demands. Those who benefit from that unfairness also do not have the standing to demand compliance with those duties if, *ex hypothesi*, those duties impose unfair demands. In such circumstances, individuals need to revise their framework. This is something they do together in a more or less deliberative process which takes reasonable account of each person's interests. They review and when appropriate revise those laws and rules which structure their pervasive interactions, whether that is in a particular institution such as a workplace, or through citizens' deliberation on the law in society at large. I shall call this their democratic obligations.

Individuals' democratic obligations express an ideal of mutual accountability through which they stand before each other to address

how their framework of co-operation will fairly accommodate their claims. Framework and democratic obligations therefore co-exist in equilibrium. Framework obligations are only genuine if they maintain genuinely fair terms of social co-operation and democratic obligations underwrite that fairness. At the same time, if democratic obligations have been adequately discharged then it is reasonable to insist that individuals meet the duties of their framework obligations, whatever burdens are involved.

It is a misconception to assume the fairness of an ideal framework of social co-operation and then insist that individuals meet their framework obligations within that. It is equally a mistake to assume that the particular, historically contingent framework actually in place in any society imposes framework duties on citizens, absent their mutual evaluation of its fairness. Adjudicatory thinking lends some support to these misconceptions because legal judgements take current laws as the baseline from which judgements are made. The appropriate baseline from which the costs and benefits of any accommodation are evaluated, is not that set of legal rules and norms which prevails at any one time because it is the fairness of that baseline which is the very thing in question.⁶ Rather, the content of the baseline needs to be informed by a normative account of what interests individuals possess, and their nature and relative strength. I shall say a little more about this as we proceed.

The notion of collective deliberative reflection on the framework of social co-operation might seem to imply the family of views known as public reason in liberal political philosophy. In the case of Rawls, the two views fit together since when citizens deliberate about basic justice (which for Rawls is essentially constitutive of the framework of social co-operation) in the political domain they must restrict themselves to public reasons. However, the notion of a democratic obligation is a broader one than that of public reason as it refers to the general ideal of interpersonal deliberation by individuals of equal moral standing. In particular, I do not want to stipulate, as Rawls and other public reason theorists do, that individuals can only propose views to each other which are grounded in shared political values rather than their particular comprehensive doctrines. One reason for avoiding that stipulative assumption is that, as Andrew March has convincingly argued, it underplays the various ways that religion *can* appropriately figure in public political debate (March 2013). Another reason is that philosophers sympathetic to religious claims such as Christopher Eberle, Paul Weithman and Kevin Vallier, have proposed an alternative convergence view of public reason where individuals may legitimately appeal to their own comprehensive religious doctrines (Weithman 2002, Eberle 2002, Vallier 2014). In what follows I take no stand on

⁶ Jones (2016) and Leiter (2013) both assume a status quo baseline, but Jones (1994) takes into account that the baseline is punctuated by cultural norms.

the debate between convergence and the orthodox ‘consensus’ views of public reason. My account of deliberative reflection on frameworks of co-operation also departs from the mainstream view of public reason in that it does not (only) apply to coercive laws, as for example Rawls assumes in restricting public reasoning to reasoning about basic justice. As a number of authors have pointed out, this is also an over-restrictive assumption (see for example Bird 2013). There are cases of for example, non-coercive establishment which can be resolved through a public reason approach (Laborde 2013). The ideal of mutual deliberation on the framework of social co-operation applies as well to individuals’ reasoning in particular domains such as employment where the rules which structure their interaction are not coercively maintained.

Rawls employs his ideal of fair co-operation between free and equal citizens in the context of principles of justice that regulate the basic structure of society. Though I agree with this picture, I want to employ the same basic ideal to describe, in idealised terms, the interactions between members of more particular institutions in a liberal society such as workplaces, universities, churches and clubs and other associations. In all of these institutions individuals interact in a way which is (a) structured by rules, and (b) realises, or frustrates, the achievement of their individual and collective aims. The rules of particular institutions are not coercive in the way the law is, though they often have legal standing, such as an agreement between employer and employed. As (a) and (b) apply in other domains besides the basic structure, it seems reasonable to import the ideal of fair co-operation into these domains too. I am happy to concede that it is the coercive law which describes the basic structure which is of ultimate importance as far as justice is concerned. But accommodation, in contrast to formal exemptions, does not typically involve formal opt outs from the law, but rather the way that laws and non-legal rules are interpreted and applied. My suggestion is that we conceptualise the individuals involved in accommodation dispute as co-members of the relevant institution who reason together on the rules affecting them all. I’ll return to this idea in Section 7; but for now we turn to consider the problems of accommodation.

3. *Objections to accommodation*

Accommodation is a controversial ideal for a number of reasons. It is said to unfairly privilege religious over other comprehensive doctrines; reward the most rigid, doctrinaire believers over those prepared to moderate their aims at some personal cost; lend credibility to grossly illiberal views such as homophobic or racist ones; and deny the fact that religious believers are partially responsible for the situation in which they find themselves, in the way that other beneficiaries of accommodation such as the disabled are not. These are all *internal* objections to religious accommodation in the sense that they are criticisms of the very idea of religious accommodation, and not objections to any

particular accommodation. In what follows, I shall try to say a little in response to all these objections, though I focus on the last which is in my view the most serious. We can distinguish these internal objections from what I shall call external costs. The latter are the particular costs, either visited on third parties by a particular accommodation or imposed upon a claimant if her accommodation request is not granted. The distinction between internal objections and external costs offers a useful way of thinking about accommodation because if we can first resolve the internal objections first, we can then employ the framework of social co-operation to assess the costs of any particular accommodation.

Let's return to the case of the Christian Sabbatarian. As I noted, her employer might reasonably say to her that she should instead take a job where Sunday working is not required, or develop an understanding of her faith where Sunday working is no longer prohibited. This is not just a theoretical view. The European Court of Human Rights (ECtHR) until recently employed its 'specific situation rule' which in effect said that the freedom to leave one's employment accorded adequate protection to citizens of faith whose religious claims could not be accommodated at work.⁷ Underlying that rule is a normative principle to the effect that individuals are responsible for their religious (and other) convictions. I shall call this *the individual responsibility objection* to accommodation. It has been advanced in different ways by a number of writers, notably Peter Jones (1994). Moreover, and no doubt partly explaining the specific situation rule, a principle of individual responsibility is assumed by Article 9(2) of the ECHR which canvasses a number of considerations (in my parlance, ways of categorising external costs) which count against an individual's Article 9(1) right to manifest her religion or belief. If we were not responsible for our religious and other convictions there would be no point in stipulating limits to how we manifest them.

The individual responsibility principle can be employed on two levels. Faced with a situation in which manifesting her beliefs comes into conflict with uniformly applicable rules, a person can either revise her beliefs or, more commonly, revise her behaviour, either by submitting herself to her employer's rules at some cost to her conscience or finding an employer which does not impose the troublesome rule. This division corresponds to two ways in which human beings are agents. As *epistemic* agents we interpret and evaluate the world we experience to form our moral, religious and other beliefs. Our beliefs do not come pre-formed and then imprint themselves upon us. As epistemic agents we are responsible for the formation of the beliefs we hold, however strongly we hold them. Holding individuals responsible as epistemic agents is consistent with their being socialised into the beliefs they regard as unshakeable convictions. The individual responsibility principle asserts only that they *could* hold alternative convictions, not that

⁷ For an analysis of the ECtHR's use of this rule, see Sandberg (2011: 84–6).

they are likely to do so or that abandoning their beliefs would not carry a considerable cost. Individuals are also *practical* agents. We pursue aims and projects which are shaped and orientated by our beliefs. It is central to our self-conception as human beings that we are active and not just passive in the world; we shape aspects of our shared world, individually and together. As with epistemic agency, this is not to deny the difficulty of revising our ends, or the costliness of pursuing ends that do not match our underlying beliefs. It is sufficient for practical agents just that we could pursue different aims than those we presently pursue.

In some circumstances, then, it clearly is possible for citizens with strong convictions to revise their practices, and possibly also their underlying beliefs. Given the external costs that accommodation can impose on third parties, it may be fairest for the state to insist on cost internalisation for believers and/or seek to engineer propitious circumstances for revision of practices or beliefs. After all, a preparedness to moderate one's beliefs and/or behaviour in the light of others' reasonable claims is a central demand of liberalism and achieving a society where citizens exhibit this moderation is seen as one of its foremost aims. The contrary policy of accommodating laws and rules around citizens' convictions, however sincerely held, has the perverse consequence of rewarding the most doctrinaire believers, over those prepared to exercise their agency to revise their beliefs and/or behaviour, thus exhibiting a central virtue of liberalism. I shall call this the *privileging the doctrinaire objection* to religious accommodation. It questions the way in which accommodation appears to reward rigidity and orthodoxy over flexibility and compromise.

Indeed, this point about doctrinaire believers can be taken further. Suppose a Christian employee worked at a hotel and had a sincerely held belief in the sinfulness of same sex relationships. When a gay couple seek to book a double room at the hotel, she refuses to give them one. In a number of jurisdictions involving cases of this kind, courts have invariably found against Christian plaintiffs.⁸ My interest, however, is in the structure of the reasoning involved in these judgements. If discrimination against same sex couples is wrong, then why should the beliefs which underlie it be accorded any weight at all in resolving the accommodation at issue? There is a difference between adjudicating in favour of a gay couple on the grounds that their interest in not being discriminated against *outweighs* a hotel proprietor's right to religious liberty—this accords the latter's discriminatory behaviour some initial normative weight—and holding to the contrary that such beliefs should not figure even as a *pro tanto* claim in any adjudication.

I shall call this the *prejudicial beliefs objection* to accommodation. At first blush, this objection may not seem a very significant one. All that's important in accommodation cases, one might argue, is that we

⁸ For example *Bull v. Hall and Preddy* [2013] UKSC 73.

get the right answer. If we reach the same destination by taking account of prejudicial beliefs as a particular kind of external cost, it surely doesn't matter too much. I believe, however, that the prejudicial beliefs objection is important, albeit for the theoretical rather than practical reason that we should not give even *pro tanto* weight to illiberal beliefs in our reasoning about accommodation. That in turn is for two reasons. First, we are considering whether rules should be interpreted and applied to take account of people's religious convictions. In contrast to the superficially similar case of hateful and offensive speech, whether that is visited on or perpetrated by religious believers, this is not simply a question of practical manifestation. It also involves the very meaning of the rules which structure our interactions. The question is why illiberal beliefs should inform the interpretation and application of those rules in even a *pro tanto* sense. Second, accommodation of the prejudicial seems relevantly similar to the issue of whether members of illiberal and harmful groups such as the Ku Klux Klan or the Mafia have special obligations to each other. It seems to me tremendously counter-intuitive to maintain that KKK members have special duties to harass or lynch black Americans or that Mafiosi owe it to each other to engage in violent criminal activity *even if those duties are over-ridden by more compelling moral considerations*. There simply are no such duties even at a *pro tanto* level. By analogy, I am suggesting, a gay couple do not have even a *prima facie* duty to accede to the prejudicial beliefs of someone who harbours anti-gay animus.⁹

As I noted, Article 9(1) of the ECHR gives individuals a qualified right to manifest their religion or belief. The ECtHR has stated that the sorts of belief appropriate for accommodation under Article 9(1) must represent 'a coherent view on fundamental problems'.¹⁰ In one UK employment case a judge accepted that the plaintiff's convictions about the need to mitigate human-made climate change was a philosophical belief of the morally right sort, and indeed the judge speculated that in the future doctrines as pacifism, vegetarianism, communism or free market capitalism might also qualify for protected status.¹¹ These are, after all, serious and important doctrines. In other recent discrimination cases in the UK, the belief that fox hunting is wrong, the spiritualist belief that it is possible to contact the dead using psychic powers, and a belief in the BBC's public service ethos were also accorded protected status (Gibson 2013: 581). The danger here is having no principled basis to draw the distinction between protected and unprotected beliefs and of opening the floodgates to a wide variety of disparate beliefs all of which would qualify for protected status. This is the *dan-*

⁹ In saying this, I am not assuming that all Christian hoteliers who refused to allow same sex couples to share a room did harbour anti-gay prejudices. I am simply making the theoretical point that if they did, then those beliefs should not figure in our deliberation about how to resolve this kind of accommodation case.

¹⁰ X v. Germany (1981) 24 D&R 137.

¹¹ Grainger plc v. Nicholson UKEAT/0219/09, paras 27–28.

ger of proliferation objection to accommodation. The objection is really the flipside to an objection sometimes made of accommodation that it unfairly privileges religious over non-religious belief. Some theorists, such as Michael McConnell and Andrew Koppelman, have responded to that objection by seeking to show how religion is uniquely special and deserving of protection (McConnell 2000, Koppelman 2006). Here I have taken a more encompassing approach: religion is special, but not uniquely so. But the basic question still remains of carving out some category of beliefs which merit special protection

Individual responsibility, privileging the doctrinaire, prejudicial beliefs and the danger of proliferation are four powerful objections to accommodation. Individual responsibility is in my view the most powerful of all as it goes to the very heart of what it is to hold a religious (or non-religious) beliefs, though nothing in what follows assumes that individual responsibility is the most important objection. At any rate, in the next two Sections I reflect on this objection further in order to show that the notion of agency which underlies individual responsibility can in fact be used to *defend* accommodation.

4. *From responsibility to identification*

A promising way to defend the individual responsibility principle is through the idea of identification. The reason we should be sceptical of accommodation claims, on this view, is because we should bear the costs of the beliefs with which we identify. The notion of beliefs needs to be interpreted in the right way so that it refers to a person's convictions, commitments, projects and so on, especially those constitutive of her identity, and not her factual beliefs. The principle which says that individuals should bear the costs of these constitutive beliefs is quite general. For example, if a person identifies as a Muslim and believes he should travel to Mecca for the *Hajj*, then we would normally think he should bear the cost of the journey. Or again, if a person strongly identifies with being an actor but is unable to get much work in the theatre, we wouldn't ordinarily think that anyone else has a duty to subsidise that project. The principle of bearing the costs of the beliefs with which one identifies (for short, the identification principle) is compatible with the idea that individuals can hardly imagine themselves not having the relevant identity-conferring beliefs. The Muslim pilgrim, if he was brought up as a Muslim from birth, may not be able to conceive of himself having any other religious beliefs, and even our none too successful actor may be unable to imagine herself doing anything else. However, the identification principle is also compatible with considering one's beliefs to be revisable; after all, individuals do sometimes change even those beliefs they consider central to their identities, for example if they undergo a profound religious conversion. The identification principle holds simply that we are epistemically competent agents who form, maintain, act on, and occasionally revise our identity-conferring

beliefs; such beliefs are not alien impositions, but rather they emanate from ourselves (even if they refer to a transcendent world beyond ourselves). We do standardly regard ourselves as responsible for our beliefs, in the sense that we are accountable for them and can be fairly criticised for holding them. We can still appropriately criticise the illiberal beliefs of a person whose racist convictions, say, are held as unshakable convictions. The aim of such criticism may only be to explain why he should not be permitted to manifest his racist beliefs; if we can get him to shift his opinions in a liberal direction, that is a bonus.

The identification principle does not, however, settle the question of what appropriate normative standard to employ in assessing how far individuals should bear the costs of their constitutive beliefs. We could for example hold that because individuals are responsible for their identity-conferring beliefs, they should bear one hundred per cent of the costs involved. That seems the most intuitively plausible judgement in the case of the Muslim travelling for a *Hajj*, for example. But in other cases our intuitions are not so clear. Should a Muslim who wants to attend Friday prayer but cannot work Friday afternoons as a result, bear the full costs of his Muslim convictions, however strongly he identifies with them, even perhaps the cost of being unable to secure full time employment (Jones 1994, 2016)? To begin to answer this, we must examine further the idea of identification at work.

5. *From identification to integrity*

I now want to suggest that the notion of identification with one's beliefs is best captured by the ideal of integrity, and that since integrity has value it grounds a *prima facie* argument *against* burden shifting, notwithstanding the cost internalisation considerations above.

The notion of integrity has been used in different ways by different philosophers. A good way into the debate over the meaning of integrity are the three categories of integrity set out by Cheshire Calhoun (1995). On what Calhoun calls the 'clean hands' conception of integrity, it consists in a person's resistance to dirtying her hands, selling out and other temptations. The person with integrity on this view 'maintains the purity of her own agency' (Calhoun 1995: 235). On the 'integrated self' conception of integrity, by contrast, the ideal consists in achieving some order and coherence between one's various aims and convictions. The person with integrity on this interpretation of the idea is not caught between incompatible aims; she marshals them into a unified whole by which she lives her life. Finally, on the 'identity view' of integrity, it consists in fidelity to those projects and principles which are constitutive of one's core identity. The person with integrity, on this third conception of the idea, ensures that her moral principles are expressed in her action and behaviour. Though Calhoun has some criticisms of them, all three conceptions of integrity are, in my view, cogent interpretations of the same general idea. Moreover, it is quite possible

for a person to exhibit more than one type of integrity through the same behaviour. For example, an artist whose over-riding aim is to be true to her art, where every other possible aim is subsumed by that, might enjoy clean hands, and identity integrity and an integrated self. The fact that, other things being equal, we would admire such an artist is a clue to the fact that on all three conceptions integrity is something that has independent value, not merely subjective value for the agent herself. At the highest level of abstraction, integrity points to the value of taking one's ideals and convictions seriously, of caring about them as only moral agents can and ought to do. Insofar as we have an interest that other human beings, not just ourselves, are moral agents, integrity understood this way has agent-neutral value.

I shall return to the clean hands and integrated self conceptions of integrity later in this Section, but I begin by focussing on the identity view of integrity since, as its name suggests, this is closest in meaning to the notion of identification sketched earlier. The identity view says that the agent with integrity expresses her commitments in her actions; the latter spring from what she most cares about, so there is not a disjunction between her outward behaviour and her inner convictions. Individuals with identity integrity live up to their convictions, not only do they not sell out (a disvalue also captured by the clean hands view), they do not sell themselves short either; they stand up for what they believe in, even if there are obstacles to doing so or theirs is a minority view. Intuitively speaking, fidelity between one's ideals and convictions and one's behaviour expresses is valuable; it is a moral virtue worth striving for (at least if one's ideals and convictions are reasonable, a point to which I shall return).

Something very much like identity integrity seems implicit in Article 9(1) of the ECHR with its distinction between belief and manifestations of belief, the assumption being that beliefs are imbricated in our behaviour, as in religious ritual for example. It is not easy to specify precisely why identity integrity has agent-neutral value, but the explanation will have something to do with the tight connection within our everyday thinking between motives and behaviour; we want to act with the right kinds of motives, this seems an important part of our well-being.¹² Identity integrity is also extrinsically valuable in the sense that it is a necessary accompaniment to other values. The person whose thoughts and actions exhibit identity integrity enjoys a certain kind of autonomy, for example, since her actions and behaviour are under the governance of her values and ideals, not anyone else's. Identity integrity is bound up with the value of self-respect; persons respect themselves for living up to their values and ideals, despite the obstacles and challenges in doing so. To be sure, this is not all that self-

¹² Bou-Habib (2006) interprets integrity as the value of a person fulfilling her subjective duties (even if they are not genuine duties) which he views as part of well-being.

respect consist of, but it is part of self-respect, and again a value that identity integrity extrinsically shares.

Individuals with identity integrity value and identify with their ideals, convictions and principles. That is why they orient their behaviour by them. If they were not *their* ideals there would be no point bearing the costs which fidelity to them entails. Identification is therefore a necessary part of integrity on the identity (and probably also the clean hands) view. After all, we would be sceptical of a person's claim to identify with a principle if she abandoned it at the first opportunity. In fact, the notion of identity integrity seems to better capture what is at stake for religious and other believers than identification alone. Religious believers do not merely endorse their religious values and consider them to be constitutive of their identities; they *value* the fact that they endorse them. Religious values are ones they prize, as evidenced in the challenges and burdens they are willing to meet in order to live up to them. Living up to one's values involves more than simply identifying with them. For example, a person might identify with the political value of conservatism, but not regard conservatism as something she is required to live up to in her personal life, nor need she even value the fact she is a conservative. Living up to one's values involves actively seeking coherence between thought and action; one's actions express one's ideals and principles, at least much of the time, and one values that expressive dimension to one's life.

Conceptualising religious commitment through the lens of integrity, not just identification, is important, because once we do so our intuitions about cost internalisation shift. I have suggested that identity integrity has agent-neutral value. That is a central value at issue in religious accommodation cases. A policy of cost internalisation, justified on the grounds that religious and other believers are responsible for bearing the costs of their beliefs, will require them to act (or refrain from acting) in a way which is at variance with their ideals and principles, often deeply held ones. This does not settle the justice of accommodation in particular instances, but it does shift the burden of justification. On the simple identification view, religious believers need to explain why others should bear the costs of their beliefs. By contrast, protecting the value of identity integrity if is a powerful consideration which *favours* accommodation despite the burden shifting it may involve.

This conclusion is bolstered if we consider the clean hands view of integrity to which the identity view is closely related. The clean hands view emphasises the ever-present temptation to succumb to social pressures and external inducements. The person who exhibits clean hands integrity preserves her agentic capacity to set her own ends for her own reasons; she does not sell out on her ends, despite third party inducements to do precisely that. She values her agency and strives to protect it from third party interference. Insofar as agency is agent-neutrally

valuable so too is clean hands integrity. After all we admire people who keep their hands clean; the person with dirty hands is corrupt, to some degree. However, in the case of religious accommodation, this claim has to be interpreted with care. Individuals who claim an accommodation are seeking to release themselves from rules which the rest of us regard as legitimate; it is those rules from which they seek to keep their agency clean. The relevant form of agent-neutral value will therefore be quite a weak one. Yet to some degree, we admire individuals who do not succumb to social pressures in the name of their principles even if those social pressures stem from laws which the majority of us regard as reasonable.

The integrated self conception of integrity articulates the value of achieving some order between one's various aims and ideals. The person with integrated self integrity enjoys some coherence between her disparate ends and pursuits, so that they are reasonably part of a unified life; she does not struggle to meet irreconcilable demands. This kind of integrity is relevant too because a fair framework of social co-operation will enable citizens to express their religious and moral convictions, and at the same time be participants in full standing in the social and economic institutions of which they are members. Individuals with strong convictions do not typically want to be exempted from society's common institutional life which after all meets many other of their interests and answers basic status needs. A person wants to be a committed Christian *and* a good employee; a dedicated Muslim *and* a keen school student, and so on. Integrated self integrity is therefore another interest which is relevant to the construction of a framework of co-operation. It too has another-regarding dimension in that it is beneficial for workmates, associates, fellow students and so on if the religious among them are able to fulfil their institutional roles and responsibilities alongside others.

6. *Three objections reconsidered*

The previous two Sections have sought to show how reflection on personal responsibility, which began as an important objection to accommodation, in fact supports it, once responsibility is conceptualised through the notion of identification and the latter explained by the ideal of integrity in its three dimensions. With integrity to hand, and also the ideal of a framework of social co-operation from Section 2, we are now in a position to address the other three objections to accommodation that I introduced in Section 3.

One of these was the prejudicial beliefs objection which questioned how theoretically satisfactory it was to accord even pro tanto weight to an accommodation claimant's illiberal views. In reply to this, it can be said that integrity only has the value it does if the commitments it involves are reasonable ones. This applies to all three senses of integrity we have been considering. With integrated self integrity, would

not for example recognise any loss of a value if a person was unable to combine his twin commitments to be a drug dealer and an armed robber. As far as identity integrity is concerned, there is only value in living up to one's moral and religious convictions if those convictions are reasonable ones, for example compatible with others' rights and liberties. I use the term 'reasonable' rather than 'valuable' because identity integrity applies also to non-moral views and also to suggest a notion of reasonable pluralism as far as people's moral and religious convictions are concerned. An individual's commitments must be reasonable in the first place for identity integrity to be relevant; we should not take the commitments which inform identity integrity at face value and then assess their reasonableness by their consistency with other values. More accurately, we should say that the value of identity integrity implies that all the components of a person's moral and religious convictions are reasonable. Thus in the example from Section 3, if a Christian hotel employee's belief that same sex couple's merit lesser civic standing which is a valueless one to live up to; not her Christian doctrine as a whole.

Moreover, the ideal of interpersonal deliberation which is at the core of the fair framework view offers a further reply to the prejudicial beliefs objection. Individuals' democratic obligations to reflect upon the rules that govern their framework involve them addressing each other as democratic partners who recognise each other's standing in determining the rules which regulate their interactions. As such, they have a duty to respect each other's co-authority in determining their common rules, and that duty is inconsistent with treating individuals with whom one is deliberating as enjoying lesser standing in all those practices which the framework governs. A refusal to offer a room to a same sex couple, conveys discriminatory message to gay and lesbian citizens, a stance which is inconsistent with regarding those citizens as participants with equal authority in a common enterprise of interpersonal justification. The religious claimants might object that they affirm that the gay couple are equally worthy human beings; they simply cannot extend a service to them in cases where (so they believe) they would be complicit in positively appraising some practice antithetical to their deepest convictions. However, this reply puts too much weight on a standalone notion of regarding someone as an equal separate from how one treats them in practice. In the shared institutional realm, how we regard people is principally manifest in the treatment we accord them. We are free to avoid others' company in the private domain and choose our friends and intimates there, but qua deliberation on common rules we have prima facie duties not to undermine each other's basic standing.

The proliferation objection questioned the possibility of circumscribing a special category of religious and moral commitments which enjoyed protected status while others do not. The notion of identity in-

tegrity offers a reply to this objection if we stipulate that it is a person's identity-related (reasonable) convictions which are at stake in accommodation cases. That seems a plausible stipulation; after all, if you do not identify with, say, being a vegetarian then it is not very coherent to claim that there is some special value in living up to your vegetarian beliefs. Adding plausibility, the notion of identity-related commitments has been employed in other areas besides religious accommodation. In his defence of minority cultural rights, Alan Patten uses the notion identity-conferring commitments to motivate a principle of fair opportunity to pursue one's cultural aims (Patten 2014: 133–136). Patten points out that identity-conferring commitments have a less negotiable character than other kinds of goals and tend to play a pivotal role in enabling a person to pursue her other goals. This way of thinking offers a response to critics of accommodation such as Richard Arneson. Arneson imagines a group of surfers who wish to take psychedelic drugs in order to transform their weekend surfing into a 'sublime and moving experience' and he questions how far this is different from religious believers who ingest hallucinogens as in the *Smith v. Oregon* case (Arneson 2010).¹³ It does not seem very plausible to say that the surfer frustrated by a law which proscribes drug-taking has failed to live up to her commitments in the same way that Al Smith argued that he was prevented from practising the elements of his Native American faith. I concede it is possible to imagine a drug-taking surfer whose commitments to those twin pursuits did form an analogous kind of identity-related commitment to Smith. But there are inevitably grey areas with any philosophical criterion and the proliferation objection only has traction if accommodation cases are peculiarly susceptible to counter-intuitive cases such as this one.

The rewarding the doctrinaire objection claimed that accommodation rewards those believers who are most rigid in their convictions and the least prepared to revise them in the light of others' legitimate claims. In a liberal society, by contrast, we should encourage individuals to reflect on their beliefs, not cultivate sectarian orthodoxies. However, while the latter is correct, it is not necessarily in conflict with accommodation. The assessment of accommodation claims proceeds from an impartial judgement on the nature and strength of relevant interests, identity integrity among them. Individuals have democratic obligations to present their views in ways that others can acknowledge and to make a genuine effort to understand others' perspectives as well. The strength of a genuine interest in some context may depart from the particular way it is presented by claimants, at least prior to mutual deliberation. The purpose of accommodation is to allow persons to live up to their reasonable identity-related commitments; this may be less than initially demanded by some claimants. More generally, the

¹³ Employment Division, Department of Human Resources of Oregon v. Smith 494 U.S. 872 (1990).

structure of interpersonal justification encourages all sides to reflect on their beliefs, in order to present them to others in ways they can reasonably accept. In some cases at least, the message of civic inclusiveness that good faith deliberation involves might promote a willingness among the orthodox to moderate their demands on others insofar as more rigid demands may stem in part from feelings of disconnection and separation from common social, economic and political institutions. It is now time to say a little more about what this deliberation involves in practice.

7. Resolving accommodation

I have set out above a framework for addressing claims to special accommodation for citizens with strong religious or moral convictions, and I have tried to show how accommodation as an ideal is not susceptible to some of the more common objections levelled against it. In this penultimate Section I outline, albeit in sketchy form, how this framework can be put to use in resolving actual accommodation cases. In doing so, I also say something about the costs which accommodation, or its absence, visits on contending parties, and how these costs might be fairly distributed.

The basic idea is that the ideal of fair terms of co-operation, as part of the larger notion of interpersonal justification, models how parties to any accommodation dispute should approach the issue at hand. Their deliberations are fair to the extent that they approach the ideal, and lacking in fairness to the extent that they depart from it. I earlier referred to individuals' framework obligations to adjust their aims so they are realisable within a given set of rules and their democratic obligations to reflect upon the fairness of the rules to which they are equally subject and to revise them where necessary. Both these sets of obligations assume a degree of responsibility, a person's responsibility for revising her aims so that they are achievable within a given framework and her responsibility to enter into good faith negotiations with others, respectively. Those negotiations will accord a high, but not absolute, value to integrity, especially identity integrity. As we noted, the participants in co-operative endeavours are citizens, but they do not for the most part discharge their framework and democratic obligations qua citizens. The notion of fair terms of co-operation is a regulatory ideal which covers other sets of rules than a state's laws; in particular it applies, or could be used to apply, to the associational domain which includes private firms as well as other employers, and educational institutions including schools and universities i.e. those domains where accommodation controversies occur. This claim needs to be interpreted with care. I am not arguing that individuals have framework and democratic obligations in every institution in which they co-operate. They do not have them in institutions not governed by codified rules, such as the family. And they need not have them in rule-bound institutions

where other values compete with fairness. For example, individuals are free to join hierarchical religious organisations the directives of which are justified first personally by religious values and third personally by the value of consent (I assume here that individuals have meaningful exit rights too). But the domains of work, and of education, are so pervasive and ubiquitous and have such a large role to play in realising, or frustrating, our interests that with very few exceptions (such as individuals who choose to join very hierarchic institutions), it seems reasonable to assert that they should approach the ideal of fair social co-operation.

To make more vivid how this ideal of accommodation could work in practice, let's return to the case of the Christian employee reluctant to work Sundays. The question for him and his employer, and his fellow employees insofar as their interests are affected, is whether his claim to be released from Sunday working should be specially accommodated. The first thing to say is that the baseline for assessing this should be a moralised one that accounts for the affected individuals' relevant interests. This contrasts with an empirical baseline where the position from which we assess whether rules should be re-interpreted or subject to exemptions simply are the prevailing legal arrangements, in this case the law which says that employers may permissibly require people to work on Sundays. It is not the case that the burden of justification falls on departures from that rule because the question at issue is whether that rule is fair in the first place and that question can only be addressed by assessing the relevant interests at stake. What are these interests? One of them is the Christian Sabbatarian's identity integrity interest in manifesting her religious convictions as well as her integrated self integrity interest in being a Christian *and* an employee in good standing and her clean hands integrity interest in not selling out her convictions. Leaving one's employment as a condition of maintaining one's religious convictions, if our Christian Sabbatarian looks for a job that does not involve Sunday working, visits upon her the cost of lost salary while she looks for work, possible retraining costs, as well as the identity, recognition, and collegial solidarity that her work gave her (notwithstanding that she may enjoy these in the future in another job). Against this, though, are the employer's and fellow employees' interests. What are those?

One of these, I think, is a general, expressive interest in having one rule for all, whether that is all employees in the case of a work organisation or all citizens in the case of a law. Uniform rules and laws signal that all count equally and are members in equal standing of the same institution, even if some employees enjoy greater responsibilities than others. This sort of expressive value counts against having separate cafeterias for managers and workers for example, or separate toilets. Of course, such 'separate but equal' policies are often discriminatory and that too counts against their implementation. But beside that,

common rules tend also to send an inclusive message that all parties are situated equally which is important for the perceived fairness of any enterprise. We can represent this as a cost, diffused among other employees, if one of their number is released from a normal requirement, such as Sunday working in the case we're considering.

The expressive value of one rule for all is a moralised interest, and there are other examples of moralised interests in accommodation claims. I have already mentioned the interests of gay and lesbian citizens in civic dignity which counts against, and in my view outweighs, the interests of Christians to be released from rules mandating their impartial treatment of individuals to whom they offer a service. Another example is the UK case of *Azmi v. Kirklees* where a Muslim teaching assistant was refused the right to wear a niqab which covered her face at work because of young children's interests in effective interaction with their teacher.¹⁴ As the latter is connected to children's moral interests in education and development it outweighed Azmi's interest in wearing her niqab at work. But besides these moralised interests we must also take account of the ordinary costs visited on an employer and a person's fellow employees, and others, if accommodation claimants have their way. Thus to accommodate a request not to work Sundays, other employees will probably have to take up the slack and re-arrange their work rota, and perhaps even work more days than they otherwise would. For another example, consider the exemptions for Sikh men in the UK and elsewhere from laws making wearing crash helmets and hard hats mandatory. If this (let us suppose) results in more head injuries for Sikhs, then this means more cases for medical staff to deal with and fractionally greater costs for taxpayers in a socialised healthcare system such as the NHS in the UK. These too are legitimate interests which must be counted against the integrity interests of male Sikh citizens.

In resolving such difficult cases, I am urging that contending sides enter into a good faith negotiation where each party acknowledges that she has framework obligations to pursue her aims within an existing set of rules but also democratic obligations to discuss and where necessary revise those rules. By referring to democratic obligations I should be clear that I am not recommending an ideal of workplace democracy where all a workplace's rules are legitimate only if they are the object of widespread deliberative agreement. I assume only that problematic rules—those that set back participants' key interests—are collectively re-assessed by those subject to them, and that a mutually satisfactory interpersonal resolution is far preferable than recourse to courts and tribunals. As far as formal exemption from a state's laws are concerned, such as the Sikh exemption from the law mandating the wearing of crash helmets, then such resolution will occur in legislatures between elected representatives, as indeed was the case when the UK Parliament passed this exemption in 1976.¹⁵

¹⁴ *Azmi v. Kirklees Metropolitan Borough Council* (2007) IRLR 434.

¹⁵ *Motor-Cycle Crash Helmets (Religious Exemption) Act 1976*, Section 2A.

8. Conclusion

The personal responsibility objection to religious accommodation, I suggested, flowed from our intuitions about individuals' identification with their beliefs. The notion of identification, I argued, cuts both ways, pointing to how individuals should bear at least some of the costs of their convictions but at the same time how their interest in identity integrity could be set back by doing so. The notion of a fair framework of social co-operation offers a way of capturing the two sets of concerns at stake. Individuals have duties both to implement the framework despite the burdens it imposes (in part because they are accountable for their moral and religious convictions), but also democratic duties to revise the framework, including re-interpreting the meaning and implementation of universal rules when they set back others' interests in integrity. Thus the notion of a framework of social co-operation offers a potentially appealing way for citizens to resolve the accommodation disputes which bedevil them in a way that fairly meets their relevant interests.

References

- Arneson, R. 2010. "Against Freedom of Conscience." *San Diego Law Review* 47: 1015–1040.
- Bou-Habib, P. 2006. "A Theory of Religious Accommodation." *Journal of Applied Philosophy* 23: 109–26.
- Bird, C. 2013. "Coercion and Public Justification." *Philosophy, Politics and Economics* 13: 189–214.
- Calhoun, C. 1995. "Standing for Something." *The Journal of Philosophy* 92: 235–60.
- Cohen, J. 2015. "Freedom of Religion, Inc. Whose Sovereignty?" *Netherlands Journal of Legal Philosophy* 3.
- Eberle, C. 2002. *Religious Conviction in Liberal Politics*. Cambridge: Cambridge University Press.
- Gibson, M. 2013. "The God "Dilution"? Religion, Discrimination and the Case for Reasonable Accommodation." *The Cambridge Law Journal* 72: 578–616.
- Jones, P. 1994. "Bearing the Consequences of Belief." *Journal of Political Philosophy* 2: 24–43.
- Jones, P. 2016. "Accommodating Religion and Shifting Burdens." *Criminal Law and Philosophy* 10: 515–36.
- Jones, P. 2017. "Religious Exemptions and Distributive Justice." In C. Laborde and A. Bardon (eds.). *Religion in Liberal Political Philosophy*. Oxford: Oxford University Press.
- Koppelman, A. 2006. "Is it Fair to Give Religion Special Treatment?" *University of Illinois Law Review* 3: 572–603.
- Laborde, C. 2013. "Political Liberalism and Religion: On separation and establishment" *Journal of Political Philosophy* 21 (2013): 67–86.
- Laborde, C. 2014. "Equal Liberty, NonEstablishment and Religious Freedom." *Legal Theory* 20: 52–77.

- Laborde, C. 2017. *Liberalism's Religion*. Cambridge: Harvard University Press.
- Leiter, B. 2013. *Why Tolerate Religion?* Princeton: Princeton University Press.
- March, A. 2013. "Rethinking Religious Reasons in Public Justification." *American Political Science Review* 107: 523–539.
- McConnell, M. 2000. "The Problem of Singling Out Religion." *DePaul Law Review* 50: 1–47.
- Patten, A. 2014. *Equal Recognition: The Moral Foundations of Minority Rights*. Princeton: Princeton University Press.
- Rawls, J. 1993. *Political Liberalism*. New York: Columbia University Press.
- Sandberg, R. 2011. *Law and Religion*. Cambridge: Cambridge University Press.
- Shorten, A. 2010. "Cultural Exemptions, Equality and Basic Interests." *Ethnicities* 10: 100–26.
- Shorten, A. 2015. "Are There Rights to Institutional Exemptions?" *Journal of Social Philosophy* 46: 243–63.
- Vallier, K. 2014. *Liberal Politics and Public Faith: Beyond Separation*. New York: Routledge.
- Vickers, L. 2016. *Religious Freedom, Religious Discrimination and the Workplace*. Oxford: Hart Publishing.
- Weithman, P. 2002. *Religion and the Obligations of Citizenship*. Cambridge: Cambridge University Press.

*Being a Progressive in Divinitia**

SEBASTIÁN RUDAS

University of São Paulo, São Paulo, Brazil

In Liberalism's Religion, Cécile Laborde defends a theory of liberal secularism that is compatible with a minimal separation of religion and politics. According to her view, liberal state—she calls it Divinitia—that symbolically establishes the historic majority's religious doctrine and inspires some of its legislation on a conservative interpretation of such religious tradition can be legitimate. In this article I analyse how is it like to belong to the minority of liberal progressive citizens in a country like Divinitia. I argue that their political activism will be defeated by Divinitia's status quo on at least four different grounds. First, in virtue of being a minority, liberal progressive citizens would rarely obtain democratic victories; second, the conservative majority could rightly argue that they do not have reasons to compromise their views in order to accommodate progressives'; third, the conservative majority can rightly complain that counter-majoritarian initiatives advanced by progressives are unfair; and four, Divinitia's public reason reproduces an asymmetry, for religiously inspired reasons can be accessible and therefore justificatory in politics, while the reasons progressives would desire to present in public deliberation would not be accessible to their conservative fellow citizens.

Keywords: Secularism, religion, church-state separation, liberalism, conservative majorities.

In June 2018, the lower house of Argentinian congress approved a bill proposing to legalize abortion in the first fourteen weeks of pregnancy. The bill was passed by 129 to 125 votes and set for discussion in the upper house on August 8. Massive demonstrations of support took place, attiring international attention and praise. On August 8, Amnesty International published in *The New York Times* a full-page message to

* Research for this article was supported by the São Paulo Research Council, FAPESP grant # 2015/12948-4. For comments and suggestions about this work, I am grateful to Ali Emre Benli, Manohar Kumar, Nicolás Rudas, Daniele Santoro, Sasikumar S. Sundaram, and Álvaro de Vita. An earlier version of this article was presented at Rijeka's Summer School Equality and Citizenship, I thank the audience for a very stimulating discussion.

Argentinian congress: “The World is Watching.” The women’s movement of several Latin American countries organized mobilizations of support to their Argentinian allies.

The bill was also strongly resisted, with the Catholic Church in the frontline of opposition. British newspaper *The Guardian* reported that Pope Francis I—or Jorge Mario Bergoglio, former archbishop of Argentina’s capital, Buenos Aires—personally requested anti-abortion senators to lobby against the bill. Under Francis I, the Catholic Church has not modified its complete opposition to abortion. The Pope’s words on it are sufficiently telling, as *The Guardian* reports:

speaking at a meeting of an Italian family association, [the Pope] said: “Last century, the whole world was scandalised by what the Nazis did to purify the race. Today, we do the same thing but with white gloves.”¹

Priests, cardinals, and bishops actively opposed the bill in their sermons. While Congress was deliberating, a “mass for life” took place at the cathedral of Buenos Aires. The upper house voted 38 against the bill to 31 in favour of it. On August 9, 2018, newspapers in Argentina and elsewhere reported the defeat of the bill as a victory of the Catholic Church.

In Argentina, a largely Catholic country, abortion is allowed only in cases of rape and when the mother’s health is in danger. Although some interpreted the result as a victory—after all, it won a majority in the lower house—there is discontent among progressive Argentinian citizens that their moral views on this particular issue cannot be decisive in political decisions. Once again, the Catholic Church showed its political muscle in Latin America, a region that counts with nearly 40% of the world’s Catholic population and which, for the first time, is the birthplace of the Pope. Progressives in Argentina complain that the Catholic Church has too much influence and power in their country, which to their eyes is a contradiction to Argentina’s public commitment to protecting liberal and democratic principles, which includes the institutional separation of church and state.

Are progressive citizens of Argentina justified in feeling politically frustrated? Is it legitimate that a historically hegemonic conservative religious tradition inspires legislation and shapes how the public sphere looks like? Cécile Laborde’s defence of ‘minimal secularism’ in *Liberalism’s Religion* (2017) offers important insights for addressing these sort of questions, as she argues that liberal conceptions of justice—to which the Argentinian state claims to be committed—are compatible with legislations that are inspired by conservative interpretations of religious doctrines. Contrary to the feeling of many Argentinian women demonstrating on August 8, the ratified restrictive law on abortion might be legitimate within a liberal state (and therefore not blatantly unjust), even if it is the case that it is inspired by a conservative interpretation of a religious doctrine.

¹ Sherwood, Harriet, “Argentina abortion defeat shows enduring power of the Catholic Church,” *The Guardian*, August 9, 2018, <https://www.theguardian.com/world/2018/aug/09/argentina-abortion-defeat-shows-enduring-power-of-catholic-church>

Progressives, both in the academic world and in society's public sphere, might be tempted to resist Laborde's position. After all, they believe their political views, although susceptible to reasonable disagreement, promote justice. According to them, their political opponents hold views that *do not* promote justice. Progressives are, therefore, in an unstable position: a common criticism in the informal public sphere directed against liberal progressive citizens is that they claim to be proud defenders of toleration and of the promotion and protection of diversity; but that when faced to practices, doctrines, attitudes, or beliefs that depart from their liberal progressive convictions, they abandon their commitments to toleration and diversity. Progressive Argentinians, the criticism could go, are committed to toleration and diversity only if abortion is not criminalized, otherwise they would not consider the restrictive legislation as just and those who defend it as reasonable. In political philosophy the charge also appears. Liberal egalitarians are criticized for having 'a tendency to frame their preferred progressive conception of justice as the only acceptable conception of justice, and to dismiss dissenters as unreasonable' (Laborde 2017: 158).²

Laborde argues that the minimal requirements of secularism in a liberal state allow for a wider range of relationships between religion and politics than liberals, or progressives,³ usually are willing to accept. In particular, she argues that religiously conservative states can be legitimate according to liberal standards of justice. Consequently, she argues that progressives should be willing to accept that, in addition to their preferred ideal worlds of progressive legislation, conservative countries that establish religion and approve legislation in accordance with a conservative interpretation of such religion can also be legitimate.

Laborde illustrates the case for a legitimate religiously conservative state by describing the (fictitious) country of *Divinitia*. This is a state that symbolically recognizes religion; finds inspiration in a religious tradition in the justification of some of its laws, restricts abortion, euthanasia, and 'other practices in bioethics; provides for religious education within the school system; grants rights of collective autonomy to religious groups in the name of freedom of association; and recognizes 'numerous exceptions and accommodations for religiously motivated behaviour' (Laborde 2017: 151–52). The legitimacy of *Divinitia* depends on whether the institutional relationship between the state and religion meets the three criteria of what Laborde calls 'minimal secularism':

- a) 'when a [religious] reason is not generally accessible, it should not be appealed to by state officials to justify state coercion;'
- b) 'when a social identity [for instance, a religious identity] is a marker of vulnerability and domination, it should not be symbolically endorsed and promoted by the state; and'
- c) 'when a practice relates to a comprehensive ethics, it should not be coercively enforced on individuals' (Laborde 2017: 150).

² Here Laborde is paraphrasing Gerald Gaus (2012).

³ I will henceforth use the terms 'liberals' and 'progressives' indistinctively.

The legitimacy of *Divinitia* situates its progressive citizens in an unfortunate position. They *subjectively* consider that certain laws and institutional arrangements of their country are unjust, yet they are required to recognize that they are *objectively* legitimate. Subjective injustice refers to the reasoned judgment of the citizen that x is unjust. In contrast, y is legitimate in virtue of its being justified according to principles are reasonable for all citizens to endorse (in this case, according to conditions a , b , and c of minimal secularism). Progressive citizens might still rightly complain that their views are not adequately taken into consideration and therefore that, even if legislation meets conditions a , b , and c , they have reasons not to accept the legislation of *Divinitia*. Laborde acknowledges this problem, and for this reason adds a fourth condition d for the legitimacy of *Divinitia*. It is a procedural one; it sets a standard of democratic fairness that consists in the inclusion of minorities ‘within fair and inclusive process of democratic deliberation’ (Laborde 2017: 156). If these conditions are met, a progressive citizen of *Divinitia* would have to put up to her bad luck of being born in a country she finds unjust—and, in cases like the Argentinian legislation about abortion, *profoundly* unjust.

The purpose of this article is to investigate what is it like to be a progressive citizen in *Divinitia*. I am interested in analysing how burdensome would it be for a progressive to live in a country that finds inspiration for some of its laws in conservative interpretations of religion. More specifically, I am interested in analysing the *prima facie* validity of transformative political activism that results from the subjective conviction of injustice. I argue that transformative political activism by progressive citizens is not *prima facie* valid once the legitimacy of *Divinitia* is accepted. In other words, the views progressives would wish to advance as reactions to their rejection of religious conservatism are rightly defeated by the conservative justifications of the laws of *Divinitia*. If it is conceded that *Divinitia* is sufficiently conservative as to trigger progressive citizens to engage in politics, then Laborde’s account of minimal secularism reproduces a *status quo* bias. The reason for this is that conditions upon which it would be adequate for a progressive to seek reform in the religiously-conservative tenets of *Divinitia*’s public sphere are very limited.

The conclusion that can be drawn from this argument is that the liberalism defended by Laborde is better suited to accommodate moderate reformism rather than progressive’s radical counter-majoritarian political activism. This is not intended to be a criticism of Laborde’s general account of liberalism. It is only an attempt to make explicit that the underlying understanding of liberalism that Laborde is presenting is in relative tension with what might be some core elements in the way some progressive citizens and movements—perhaps including some variants of the feminist movement—consider their main political motivations.

1. *On how conservative is Divinitia*

Although Laborde claims that a country that finds inspiration for some of its legislation in a conservative interpretation of a religious doctrine can be legitimate according to a standard conception of justice, she does not make explicit references as to *how* conservative such legislation can be. She says, for instance, that in *Divinitia* ‘there are restrictive laws about abortion’ (Laborde 2017: 151), yet she does not specify how restrictive these laws can be. She also suggests that the establishment of the Anglican Church in Great Britain could be understood as conferring Britain some of the features of *Divinitia*, yet she stresses that British legislation about abortion and same-sex marriage is far from being conservative—with the exception of Northern Ireland (Laborde 2017: 153). How conservative can *Divinitia* be?

Religious conservatism in many western democracies is salient in issues that are directly and indirectly related to regulations of sexuality and sexual behaviour. Thus, a conservative interpretation of the religious doctrine of the majority might inspire legislation on abortion, family law, education about sexuality in elementary school, and the definition of the family as heteronomous and monogamous. In recent years, governmental initiatives of mandatory HPV vaccination received vigorous opposition by both vaccination sceptics and religiously conservative groups, the latter most probably because of the fact that HPV is a sexually transmitted disease.⁴ What are the limits that these legislations must observe in *Divinitia*? An immediate constraint to *Divinitia*’s legislation is that its basic structure is regulated by a liberal conception of justice. This means that it is committed to protecting the core set of basic liberal rights such as freedom of conscience, freedom of association, equality before the law, rule of law and equal political rights. *Divinitia* is conservative, but within the limits of liberalism. What are these? A couple of examples can help to illustrate.

Opposition to same-sex marriage is one of the typical positions that conservative religions uphold in western democracies. Can *Divinitia* have a conservative religiously inspired regulation on marriage? Considering that *Divinitia* endorses a liberal conception of justice, its regulations on marriage cannot go against the idea of citizens as free and equal. For instance, there would not be a ban on homosexual relationships, even if some conservative interpretations of the Bible would suggest otherwise. If what is at stake is free and equal citizenship, conservative religion cannot inform legislation.

Yet not all of conservative religion on issues of same sex legislation interferes with free and equal citizenship. Legislation on same sex marriage can be taken to the symbolic realm. Some religiously inspired opponents to state’s recognition of same-sex marriage argue that ‘mar-

⁴ Grimes, David Robert, “We know it’s effective. So why is there opposition to the HPV vaccine?”, *The Guardian*, January 11, 2016, <https://www.theguardian.com/science/blog/2016/jan/11/why-is-there-opposition-hpv-vaccine-cervical-cancer>

riage' is by definition the union of a man and a woman, and that calling 'marriage' any other form of union (a same sex couple, for instance) is a conceptual impossibility. This form of opposing to same sex marriage can be accompanied with the proposal to recognize civil partnership for same sex couples. Thus, same sex unions would be recognized by the state, yet they should not be called—and registered—as marriages. In practice, a same sex union would receive the same rights than a marriage receives, yet they would be named differently.⁵ The inequality that is institutionalized is only symbolic. Can this be a feature of *Divinitia*?

One immediate response would point that this is not a mere *symbolic* unequal recognition, for it entails discriminatory behaviour against those who are in 'civil unions.' Non-recognition of same sex (amorous) relationships as marriage creates an unnecessary instance in which LGBTI members are obliged to disclose a source of discrimination against them. This can happen, for instance, at the moment of filling out registration forms that ask about marital status. This does not affect all equally, but creates situations in which unwanted release of information becomes mandatory and disadvantageous for members of a group that is already in condition of vulnerability. This argument against unequal symbolic recognition shows that the inequality at stake is not entirely symbolic, for it triggers discriminatory behaviour. Therefore, its legitimacy might be contested in virtue of making more salient vulnerable social identities (condition *b* of minimal secularism).

It is possible to think of a case in which this discriminatory behaviour is not triggered and therefore unequal symbolic recognition remains entirely symbolic. Let's assume that society in *Divinitia* does not discriminate against homosexuals, yet it still defends the unequal symbolic recognition—in virtue of a conservative interpretation of a religious doctrine. Under this assumption, unequal recognition remains entirely symbolic as same sex unions and marriage unions would be granted equal rights and no social discrimination would occur against individuals who constitute the former kind of union.⁶ Following the criteria of minimal secularism, in this case unequal symbolic recognition is legitimate if the justification provided for this unequal symbolic recognition is advanced in accessible public reasons (for instance, on the etymology of the word 'marriage,' on its traditional and historic meaning); if it does not deepen social vulnerabilities (that is, that it does not harm LGBTI members in any sense similar to the one described above); if it does not entail that all citizens must accept the truth of a religious doctrine (the regulation is inspired by the religious doctrine, yet it does not impose it); and if fair democratic procedures are ob-

⁵ Notice that, in this example, civil unions would recognize the right to adopt children to same sex couples. Religiously conservatives are usually against such a recognition.

⁶ For a critical analysis of whether purely religious reasons for opposing same sex marriage can be accepted by Laborde's conception of public reason, see (Bardon 2018).

served in the process of its approval. Under these conditions, this form of differentiated recognition can be legitimate. Since disagreement is a permanent condition of democratic regimes, it is normal to assume that there will be reasonable citizens that would deeply disagree with regulation of this sort. They might indeed find it offensive and unjust, and they might have good reasons to support their beliefs. These citizens must however recognize the legitimacy of the regulation and accept its authoritative status.

The second case is abortion. In *Divinitia*, Laborde says, there will be restrictive laws about it. How restrictive these laws can be? This is a harder question to address because disagreements about the permissibility of abortion are often described as an illustration of a debate that leads to reasonable disagreements. In a reasonable disagreement, both parties are epistemically justified in their beliefs, even if these lead to opposite conclusions. Reasonable disagreements are originated by cultural, epistemic, and idiosyncratic reasons—among others. They are particularly challenging in politics because of the urgent need to make decisions, and in cases of reasonable disagreements some sacrifice needs to be done. If a particular issue is a case of reasonable disagreement, and if conditions *a*, *b*, and *c*, are observed, then it seems that the legitimacy of the regulation to be approved would depend on the democratic fairness of the decision. Here democratic majorities are advantaged and therefore a conservative or restrictive legislation—in the case of abortion—would be approved. But how conservative—or restrictive—this legislation can be? What are the limits set by liberal justice in this specific issue?

A typical conservative (and religiously conservative) argument against abortion appeals to the value of protecting (innocent) human lives from being killed. As such, this is a reason that any reasonable person is expected to accept. The source of the disagreement in debates about abortion lies in the difficulty to determine whether abortion involves ‘killing innocent human lives.’ Here the debate turns into a question about the beginning of human life, and on this issue there is opacity: at some point in pregnancy the foetus turns into a human being. This opacity makes the conservative argument a potentially valid one—although not necessarily an unbeatable one. As long as this specific question is not answered beyond any reasonable doubt, it is impossible to reject the conservative’s view as unreasonable.

Although there is opacity, it is possible to identify some limits to the degree of opposition to abortion by the conservative, which means that various of the most radical religiously inspired views on abortion would not be permissible in *Divinitia*. As in the same sex marriage debate, liberal justice sets limits. First, concern for human life might favour the life of the mother when her life is in serious risk as a consequence of the pregnancy. Second, if it is established that the foetus suffers from severe malformations that make its life outside of the uterus non-

viable, then it might be unreasonable to oppose abortion. Third, bans on emergency contraceptives would be ruled out. Reasonable disagreement about the beginning of human life is not about whether human life starts right away after conception—or within the first five days, which is the usual timeframe of efficacy of emergency contraceptives. Here the disagreement is about where to draw the line that indicates the beginning of a *human* life. Today, however, there is consensus that this does not happen right after conception—the main exception among the most vocal actors in this debate being the Catholic Church, which, as a matter of fact, did not embrace this view for the most part of its history—,⁷ but somewhere after the first trimester, which explains why much debate about legislation on decriminalization of abortion usually sets a limit that avoids late abortions. This indicates that emergency contraceptives would not be criminalized in *Divinitia*.⁸ Against very strong convictions by feminists and other progressive citizens, restrictive legislation on abortion based on the indeterminateness about the beginning of human life would, in principle, include the restriction on abortions in cases of pregnancies resulting from sexual violence. Liberal justice, however, *might* require exceptions to this restriction. For instance, in cases in which the victim of sexual violence is underage. Other values protecting the inviolability of human dignity might override the grounds leading to the prohibition in cases of rape, yet it might be the case that the reasons leading *Divinitia* to legitimately approve restrictive legislation on abortion apply to cases of rape as well.

Divinitia's legitimacy in enacting restrictive legislation about abortion, therefore, does not imply that it could be banned altogether. It is most likely that restrictive legislation on abortion would refer only to

⁷ Dworkin argues that in public debate people do not really define their position on abortion in relation to the question of the beginning of life. Instead, he argues, what concerns people the most is an idea about the sacredness of life. In the case of the official view of the Catholic Church, which opposes any form of abortion on the grounds that human life starts at conception, he says that average Catholics do not believe in such views. Many conservative Catholics accept that abortion should be permitted *in some cases*. According to Dworkin, this shows that they consider that, although human life is very important (sacred), there are overriding reasons that lead them to accept it could be sacrificed. Therefore, their opposition cannot be that abortion should not be permitted because it involves the sacrifice of a human life (Dworkin 2011: 39).

⁸ According to the International Consortium for Emergency Contraception, 147 countries have at least one emergency contraception (EC) pill brand registered, 95 of which allow access to EC without prescription, which means free distribution of EC. 60 countries are registered as including EC in their essential medicine lists. 47 countries have no emergency contraceptive pill brands registered, which amounts to a complete restrictive policy. In Argentina, EC is not included in the Essential Medicines List (as of 2005). Access to EC requires medical prescription, the distribution of which is regulated within post-rape care guidelines. Source: International Consortium for Emergency Contraception <http://www.cecinfo.org/country-by-country-information/status-availability-database/> (last accessed September 3, 2018).

cases that relate to the voluntary termination of unwanted (and, controversially, probably also forced) and advanced pregnancies of adult women, when their lives are not under risk. If Argentina were a well ordered society, its current legislation on abortion would be legitimate in spite of its recent multitudinous demonstrations in favour of decriminalization by the feminist movement (and a great number of its supporters that might not describe themselves as feminists). The feminist movement, however, happens to be in the minority and therefore loses to conservatives. In the next section I investigate whether counter-majoritarian political strategies—the only ones that might seem feasible for a political minority such as the feminist movement—are *prima facie* valid in a country like *Divinitia*.

2. Progressives and the status quo in *Divinitia*

In this section I show that minimal secularism reproduces a *status quo* bias. By this I mean that it justifies the preservation of the *status quo* by making it harder for reformers to advance their transformative projects. I argue that if it is accepted that *Divinitia* can be legitimately conservative, then the conservative majority is morally entitled to preserve the conservative tenets of the public sphere and legislation of its country. Although in *Divinitia* progressive citizens can access procedural democratic mechanisms to transform their society (this is condition *d* of minimal secularism), I claim that they will face a valid opposition grounded on the plausible view that they should respect the legitimate *status quo* of their society.

Divinitia adequately justifies its laws, is inclusive, does not impose a comprehensive ethics, and observes inclusive processes of democratic deliberation. Importantly, progressive citizens of *Divinitia* will consider that some key aspects of the legislation of their country are profoundly unjust. However, given that the country is sufficiently democratic, they have to accept their legitimacy and obey the law in spite of their discontent. Notice that this is not a minor requirement, for it means to put up with legislation that they consider deeply unjust. Although progressive citizens can be politically engaged in order to promote social, political, and legislative change, I will now argue that they will have to struggle with a *status quo* that rightly justifies itself.

Before presenting the *status quo* bias argument, it is important to address a potential objection. It could be argued that the requirement to put up with legislation one disagrees with is a normal feature of a healthy democracy. Indeed, that accepting such kind of requirements is what makes democracy possible. It is not my interest to contest this fundamental feature of democracy. I think, however, that the case at hand in *Divinitia* is different. The population I am referring to as progressives are those who do not embrace the conservative interpretation of the religion that inspires part of the legislation—particularly the one related to the family, sexuality, and reproduction—of *Divinitia*. The

requirement is therefore not only to accept the fact that it is possible to lose in democratic contestation, but that one is losing to a political view that is inspired by a conservative interpretation of the religion embraced by the majority. In other words, to accept that one is being defeated against a majority that promotes its conception of the world in politics. Although restrictive legislation on abortion or on same sex marriage is justified in terms that it respects minimal secularism, it remains true that such legislation is approved *because* it does not contradict the conservative interpretation of the religion embraced by the majority. The requirement is not therefore merely to accept democratic outcomes, but to accept that legislation in matters one considers extremely important is decided in relation to an interpretation of a religious doctrine one deeply disagrees with.⁹ To this requirement it must be added that progressive citizens would be required to accept that their political activism, although legitimate, can be defeated by the conservatives' will to preserve the *status quo*.

The *status quo bias* argument can be illustrated by appealing to David Miller's argument of 'historic precedence.' According to him, the territorial historic majority of a country is entitled to privilege in the public sphere and in the legislation to its cultural and religious identity. Discussing the referendum that proposed to ban the building of minarets in Switzerland, Miller defends the claim that cultural majorities are entitled to ensure that the appearance of public space reflects its own cultural values. As he puts it, 'if such values reflect a Christian heritage, then a Christian public sphere and legislation might remain hegemonic' (Miller 2016: 448). According to him, a people that transform and occupy a territory according to its needs creates a value the enjoyment to which it is entitled to. Part of this value is symbolic, 'as the territory comes to bear the imprint of the national culture' (Miller 2016: 448). Members of the cultural majority understand their historic identity 'partly through their direct experience of the environment they and their predecessors have created' (Miller 2016: 448). The value created by this historic process, Miller concludes, is 'the value of national identity' (Miller 2016: 448). The entitlement to preserve this historically built national identity held by the indigenous majority is what can justify privileging the cultural identity of the majority. Restrictions on cultural or religious expressions of immigrant minorities might be justified by appeals to the entitlement of preservation and privilege of the majority's cultural identity. In the case of the Swiss referendum, this could mean a ban on the building of Islamic minarets for the sake of protecting (Christian) national identity.¹⁰

⁹ What is at stake here is that it is religion what inspires legislation. I am assuming that, whatever its further meaning, in western cases of *Divinitia* (i.e. Argentina), religion concerns existential questions. This understanding of religion is not unusual among political philosophers (Nussbaum 2009: 168; Maclure and Taylor 2011: 12–13).

¹⁰ It is important to annotate that Miller's conclusion does not vindicate the

Minimal secularism provides a straightforward answer to Miller's argument—if used to justify the Swiss ban on the building of minarets in Muslim mosques (Laborde 2017: 138–39). If the national identity of the country is defined in religious terms (as it is the case in the Swiss ban), and if religious identities are markers of vulnerability and domination (as it is also the case in the Swiss case), then privileging a Christian heritage in the name of national identity—even if this is done by attributing to Christianity a cultural, not religious, nature—is not legitimate according to the liberal standards of minimal secularism. The ban is illegitimate because it creates first-class and second-class categories in the definition of Swiss national identity. This categorization is generated by the facts that national identity is defined in terms of religious heritage and that grouping individuals alongside categories of religion is a form of reinforcing social vulnerability of some individuals—that is, the Muslim Swiss minority.¹¹

It is important to notice that the objection to Miller's argument shows that the ban on minarets is illegitimate in the specific case of the Swiss referendum. It does not imply that a ban of such kind will always be illegitimate. If context shows that conditions *a*, *b*, *c* and *d* are met with the ban, then minimal secularism would allow it. If conditions are met, minimal secularism can be compatible with bans on Muslim minarets. Minimal secularism also allows for a majority to privilege its religious or cultural heritage. This is what happens in *Divinitia*. In other words, not all privileging of a conservative, traditional, and religiously inspired culture is illegitimate. Similarly, not all banning of certain religious expressions is illegitimate—although the mere discussion on a ban of this sort might already be an indicator that, for instance, religion is a marker of vulnerable social identities. In a context where this is the case, what can a progressive do in order to transform her society in such a way that it adjusts its institutions and legislation in accordance to what she believes is just?

It is possible to identify two alternatives open to progressive citizens. The first one is to engage in political activism participating in democratic politics. By doing so progressive citizens might eventually succeed in changing what they think needs to be changed. This, however, is a frustrating task, because progressives are a democratic mi-

Swiss referendum entirely. His argument is that it is permissible that a historical cultural majority privileges its cultural heritage by restricting cultural expressions of the cultures of immigrant minorities. This means that restrictions on the building of minarets in Switzerland are permissible. He, however, advances two criticisms to the Swiss referendum: its outcome should not have been entrenched in the constitution and it should have been held at the canton-level and not at the national level (Miller 2016: 452–54).

¹¹ Religious identities are not always markers of social vulnerability. In order to show how this is so, Laborde refers to the cases of Senegal and Madagascar (Laborde 2017: 142). One might wonder whether these are the two exceptions that confirm the rule, thereby making minimal secularism's criterium on vulnerable identities extremely demanding.

nority which therefore might not be able to obtain enough support to transform legislation or institutional arrangements. Progressive citizens can no doubt benefit by the democratic fairness that is in place in a country like *Divinitia*. This ensures political rights and the possibility to actually form political movements that could aspire to democratic influence when the time comes.

A second alternative is to appeal to strategically counter-majoritarian political engagement (for instance, judicial activism). This form of democratic activism consists of addressing the judicial branch of the state in order to press it to promote legislative changes. This counter-majoritarian political engagement has proved efficacious in legislation transformation in several countries; for instance, activism for decriminalization of abortion has taken this strategy in some countries where Catholicism is the historic majority religion.¹² Although this is probably an attractive and promising strategy in terms of efficacy, there are two reasons that challenge it. According to the first, the conservative majority does not have reasons to compromise the *status quo*. According to the second, the conservative majority would be right in arguing that any change would be an unfair imposition upon them. Let's consider each one separately.

Conservatives in *Divinitia* could rightly reject progressives' unreasonable views. For instance, if they were attempting to impose atheism as the state's doctrine. This is not, however, the kind of political initiative I am attributing to progressives of *Divinitia*. I am assuming that they are advancing initiatives that are legitimate according to the framework given by minimal secularism. These are initiatives that could easily be approved in *Divinitia's* neighbour *Secularia*.¹³ In other words, they are legitimate views that, if held by a democratic majority, should be accepted—conservative citizens of *Secularia* would have a duty to accept them even if they are subjectively unjust. However, in *Divinitia* the majority accepts conservative legislation and therefore opposes to the initiatives progressive citizens try to advance. This is not capricious opposition grounded on political sectarianism. Conservatives in *Divinitia*, that is the majority, (rightly) claim that conservative legislation is just and that it has been legitimately approved. They will rightly claim that they do not have reasons to be open to modify the law, because the law respects the three criteria of minimal secularism and they have observed democratic fairness in decision-making processes. Although they can recognize the political reasonableness of the initiatives their fellow progressive citizens advance, they think of them

¹² For instance, in Colombia, where abortion is decriminalized only if pregnancy is the outcome of rape, the life of the new-born is at serious risk due to serious foetal malformations, and the life of the woman is in serious risk due to pregnancy-related complications.

¹³ *Secularia* is the mirror country of *Divinitia*, that is, it is a fictional legitimate country the legislation and public sphere of which are largely inspired by secular and progressive ideals (Laborde 2017: 151).

as unjust and ultimately unacceptable for their society. The conservative majority, therefore, would be right in claiming they do not have any reason to incorporate the views of the progressive minority.

Additionally, conservatives of *Divinitia* could also complain that it would be unfair against them if they were forced to live under a political regime that does not reflect their conservative views in the public sphere. The conservative majority has collectively shaped the public sphere of its society in such a way that, first, reflects their beliefs and traditions and, second, succeeds in giving due respect to its minorities. Hence it would be unfair to suddenly change the public sphere in such a way that does not reflect the beliefs and practices of the conservative majority and that would consider the new legislation as (subjectively) unjust—yet probably (objectively) legitimate. Progressive citizens of *Divinitia* are therefore fated to accept their bad luck of being in *Divinitia* and not in their preferred *Secularia*.

A consequence of the *status quo* bias is that some famous progressive counter-majoritarian victories lose legitimacy precisely in virtue of their counter-majoritarian nature. *Roe vs. Wade*—the 1973 United States' Supreme Court decision that recognized women's right to decide to have an abortion during the first trimester of pregnancy—has been described as having introduced the recognition of a right on a very contested issue that, back in the 70s, might not have enjoyed widespread popular support. The decision of the Court was introduced against a conservative—and probably legitimate—legislation and, most likely, against the conservative social ethos of the population. Even if it is accepted that the decision of the Court advances a legitimate goal, it is also true that it disrupted the legitimate expectation of the conservative majority to live in a country that reflects its social ethos.

Counter-majoritarian decisions are sometimes criticised for being counterproductive. While they might be well intentioned and social justice promoters, the criticism goes, their abruptness generates division and polarization that contributes to perpetuating conflict rather than overcoming it. In other words, in the debate about decriminalization of abortion in the U.S., they suggest that had the Court handled the abortion issue on a step by step basis, abortion would not be such a salient social cleavage in contemporary politics, for it would have been already overcome. Following this line of thought, the adequate form of political activism would be to pursue gradual reforms instead of counter-majoritarian measures. To this *prudential* argument in favour of moderate progressive reformism and against progressive radical counter-majoritarian political engagement it is now possible to add a *principled* one, which defends that, in *Divinitia*, counter majoritarian radical political engagement is likely to be defensible and legitimately so.

An alternative for the political engagement of the progressive citizen of *Divinitia* could be political resistance. By it I mean that progressives' political engagement could not be aimed at transforming legis-

lation but at impeding it from being approved. Given that they are a minority, this form of political engagement might likely be bound to fail. A possibility that is open to progressives in *Divinitia* is to function as watchdogs and control the legitimacy of legislation that is inspired by conservative interpretations of religious doctrines. This strategy is fairly common in the public sphere when public debates turn to abortion or same sex marriage. Progressives usually claim that conservative views are religiously inspired and that the separation of church and state proclaimed in liberal democracies impedes them.

The interpretation that Laborde offers about public reason shows that progressives are wrong in thinking that no religious inspired argument can function adequately as public justification. She maintains that public reasons should be accessible. A reason is accessible when it ‘can be understood and assessed, but need not be endorsed according to common standards’ (Laborde 2017: 120). If ‘actual (not idealized) publics’ find a reason accessible, then such reason is an adequate reason for public justification. Public reason, then, provides the ‘conditions of possibility of public debate’ (Laborde 2017: 120–21) in so far as members of the public are able to understand each other even if they disagree in the substantive content each one advances. Laborde says that her theory of public reason is empirical (Laborde 2017: 128), which means that its determination of the set of public reasons of a particular society, for instance *Divinitia*, would depend upon *Divinitia’s* specific characteristics: how much of an open society *Divinitia* is will highly influence its set of public reasons. As she puts it ‘in pervasively religious communities, religious reasons *strictu sensu* [...] may well provide the only currency of public reason.’ This is so because religious reasons ‘provide a common currency of argument and debate’ (Laborde 2017: 128).

Laborde is right in pointing out that it is highly unlikely that any society has such degree of homogeneity—if that is the case, it might be because of the ‘oppressive use of political power,’ as Rawls famously put it (Rawls 2005: 37). The possibility that a religious doctrine is the currency of public reason shows, however, that argument and public debate in *Divinitia* can be highly influenced by *Divinitia’s* shared public culture. This culture would be inspired by a conservative interpretation of the historic religious doctrine, the key facets of which would be well known and understood by citizens of *Divinitia*. Public education would make sure that everybody gets some understanding of *Divinitia’s* major religious tradition; public holidays will commemorate important events for such religion; legislation will be inspired by religion, and the public sphere will give prominence to the historic religious tradition. In a context like this, it will be only natural that a set of religious reasons will be accessible to any citizen of *Divinitia*, progressive or not.

How does this conception of public reason impact progressives’ political resistance? Given that they have been socialized in *Divinitia* and its institutions, progressive citizens would not be able to claim that

they do not understand, or that they cannot engage with, justifications that appeal to conservative interpretations of religious doctrines, or, for that matter, that include religious references in their justifications. For them, it would not be true that they cannot access such reasons, for they have been raised in a culture which makes sure certain reasons are accessible to anyone. Progressive citizens are dissenters in their society in spite of the fact that conservative religiously inspired reasons are accessible to them.

The kind of dissent progressive citizens espouse originated in the fact that they remain unconvinced by the moral appeal of the reasons (accessible to them) that support part of the legislation of *Divinitia*. Here they face a problem, because it is *their* reasons which will not be adequately socialized in *Divinitia's* public culture. The lack of socialization of their reasons makes them inaccessible to their fellow non-progressive citizens, who might not be able to understand the appeal of feminist-inspired views on why abortion should not be criminalized; or why queer theory-inspired views on gender and sexuality should be taken into consideration the moment legislation on the institution of the family is approved—assuming, of course, that these are the sources of inspiration of the views supporting both abortion and reform in regulations about the institution of the family.

I have argued that the *status quo bias* that *Divinitia* reproduces affects progressive citizens in four different ways. First, the likelihood of gaining democratic majorities is reduced and therefore their political claims will be systematically outvoted. Second, minimal secularism entails that the religious and cultural majority of *Divinitia* can be entitled to preserve its national culture, which implies that they do not have reasons to compromise their views in order to accommodate the political claims of the minority of citizens that constitute the progressive side. Third, from the fact that progressives are constantly outvoted it does not follow that conservatives have a *prima facie* reason to attend their claims. A requirement to do so would be unfair. Fourth, progressive citizens will have limited tools to enter in public deliberation because their fellow citizens might not be able to understand and assess the reasons they would like to advance in order to support their political claims. Being a progressive in *Divinitia* entails the acceptance of a very limited space for hope in the possibility of political, let alone radical, transformation.

3. Conclusion

The purpose of this article has been to explore how is it like to be a progressive in *Divinitia*. Contrary to what it might be initially thought, liberal justice sets clear limits to what can be approved as legislation in *Divinitia*. Legislation about abortion might be determined in terms of the (religiously inspired) question about the beginning of life. However, the current official doctrine of the Catholic Church—that uncon-

ditioned protection of human life, which begins at the moment of conception, should be guaranteed by the state—would not be legitimately approved. Legislation about same-sex marriage might also be defined in relation to the defence of traditional, religiously inspired conceptions of the family. However, liberal justice requires that members of LGBTI population enjoy of the same package of rights that any other citizen enjoys. Again, some of the most conservative oppositions to same sex marriage and relationships would not be accepted in conservative yet minimally liberal *Divinitia*. Although these limitations could be seen as creating a tolerable political environment for *Divinitia*'s progressive citizens, in the second part of this article I have shown that minimal secularism is biased towards the preservation of the *status quo*. I argued that the legitimacy of *Divinitia* means that progressive political advocacy loses currency and legitimacy. The political strategies open to progressive citizens for transforming what they think is unacceptable religious conservatism are likely defeated by arguments that defend the *status quo*. Against counter-majoritarian strategies, the conservative majority can reply, first, that they do not have any reason to compromise the views that inform how legislation is shaped, and second, that a requirement to compromise would be unfair. Against what I called progressives' political resistance in public deliberation, the conservative majority can reply that there is an asymmetry between the adequacy of the reasons they offer to progressive citizens in order to justify legislation that is inspired by a conservative interpretation of a religious doctrine, and the adequacy of the reasons progressive citizens offer in order to transform such legislation. In *Divinitia*, the former are likely to be accessible to all citizens, while the latter are not.

As mentioned at the beginning of the article, I have not intended to advance a criticism to Laborde's conception of minimal secularism. I have tried to specify the consequences of accepting the legitimacy of a country like *Divinitia*. Although *Divinitia* would be entitled to inspire its legislation on a conservative interpretation of its hegemonic religious tradition, its institutions will be constrained by a conception of justice that is distinctively liberal. In this article I have made apparent that the liberalism of this conception of justice is of the type that prefers moderation in political activism over radical political activism advocated by some contemporary social movements.

References

- Bardon, A. 2018. "Is Epistemic Accessibility Enough? Same-Sex Marriage, Tradition, and the Bible." *Critical Review of International Social and Political Philosophy* 0 (0): 1–15. <https://doi.org/10.1080/13698230.2018.1487231>.
- Dworkin, R. 2011. *Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom*. New York: Knopf Doubleday Publishing Group.

- Gaus, G. 2012. "Sectarianism Without Perfection? Quong's Political Liberalism." *Philosophy and Public Issues (New Series)* 2 (1): 7–15.
- Laborde, C. 2017. *Liberalism's Religion*. Cambridge: Harvard University Press.
- Maclure, J. and Taylor, C. 2011. *Secularism and Freedom of Conscience*. Translated by Jane Marie Todd. Cambridge: Harvard University Press.
- Miller, D. 2016. "Majorities and Minarets: Religious Freedom and Public Space." *British Journal of Political Science* 46 (2): 437–56. <https://doi.org/10.1017/S0007123414000131>.
- Nussbaum, M. 2009. *Liberty of Conscience: In Defense of America's Tradition of Religious Equality*. New York: Basic Books.
- Rawls, John. 2005. *Political Liberalism: Expanded Edition*. 2nd ed. New York: Columbia University Press.

Laborde's Liberalism's Religion: The Problem of Religious Exemptions

MIKLOS ZALA
Central European University, Budapest, Hungary

In this paper, I critically examine Cécile Laborde's Liberalism's Religion and argue that her approach to religious exemptions faces significant difficulties. I first highlight some methodological disagreements with Laborde's theory. I raise concerns about her theory's 'two-pronged' structure being too narrow. Moreover, Laborde's 'disaggregation approach' promises a context-sensitive, bottom-up theory of exemptions which examines exemption claims on a case-by-case basis, but instead offers a top-down theory that provides an idealized explanation for potentially all religious exemption cases. I argue that a non-ideal approach which does not offer an overarching explanation of exemptions is preferable to Laborde's. Next, I discuss further problems with Laborde's theory, which concern her assumption that if there is something 'ethically salient' about religious practices, it must be located at the personal level. Laborde claims that if we want to ascertain the ethical salience of a practice, we must focus on the relationship between the person and her commitments. But this individualistic focus cannot always account for why we want to accommodate religious practices. Such practices, I claim, are sometimes accommodated not on an individual, but on a group-based rationale. Finally, I address Laborde's dismissal of the analogy between religion and disability. Laborde's view regarding disabilities and the stated analogy is unsatisfactory in two respects: it is based on the medical model of disability and it overlooks the role of the environment in turning physical impairments into disabilities.

Keywords: Liberalism and religion, religious accommodation, religious exemptions, multiculturalism, disability accommodation.

1. *Introduction*

The relationship between religion and politics has fascinated political and legal theorists for a long time. Since most liberals hold the ideal of state neutrality in high regard, special accommodation of religion in

law is not easily justified. One question that liberals are particularly interested in is whether exempting religious practices from generally applicable laws is justifiable without giving religions unjust privilege. The uneasiness with religious accommodation is not difficult to understand: although religion's prominent role in society has weakened, religious exemptions permit members of certain religious groups freedoms that non-members do not enjoy (cf. Levy 1997: 28). This is indeed puzzling for liberals who believe the state should not favor one conception of the good over others.

In her new book, *Liberalism's Religion*, Cécile Laborde offers a novel framework for reconsidering the question of religious exemptions.¹ Her alternative view to earlier, mainstream approaches that try to analogize religion with other 'conceptions of the good' is the 'disaggregation approach,' a concept that she adopts from the work of James Nickel (see Laborde 2015: 594). In Laborde's understanding, "[t]he starting point of the disaggregation strategy is to suggest that different parts of the law should capture different dimensions of religion for the protection of different normative values" (Laborde 2015: 594). The view posits that we should not look at religion through a single lens. Rather, we ought to analyze what religion *is* in the given context and what values are at stake. Thus, throughout the book, Laborde devotes great effort into the examination of what 'religion' in the given context is, and what kind of normative response it prompts.

I am convinced that *Liberalism's Religion* will become an instant classic. But despite its elegance, ingenuity and importance, I will argue in this paper that Laborde's approach to religious exemptions faces serious difficulties. I will contrast her approach with my earlier work that I believe provides more convincing solutions to the highlighted challenges.

This article proceeds as follows. Section 1 provides a brief outline of Laborde's theory of exemptions. The three subsections of Section 2 flesh out my three distinctive disagreements with Laborde's approach. The final section concludes.

2. Laborde's theory of religious exemptions

In *Liberalism's Religion*, Laborde rejects three promising liberal strategies for solving the puzzle of religious exemptions—*dissolving* religion, *mainstreaming* religion and *narrowing* religion. The strategy of 'dissolving religion' acknowledges that it is indeed unjustified to exempt religious practices from the requirements of generally applicable laws. This is Ronald Dworkin's position, which holds that religion and religious freedom are not themselves special (2013). Thus, Dworkin's answer to what Laborde calls religion's 'ethical salience problem', i.e.

¹ The other big question her book addresses, which I do not discuss in this paper, is religious establishment.

the evaluation of the ethical relevance of “different kinds of beliefs, practices, and identities” (Laborde 2017: 41) is that religious belief is just one type of comparable non-religious commitments, worldviews, and (even seemingly mundane) preferences (Laborde 2017: 44). This is why Laborde labels Dworkin’s stance on religion as ‘dissolution’:

[Dworkin’s strategy] proposes to broaden religion into a maximally inclusive category that comprises preferences, commitments, identities, beliefs, worldviews, and so forth. Religion is not so much analogized with them as dissolved into them. This means that neither religion nor the more inclusive category it falls under needs to be defined precisely, because their boundaries and scope are irrelevant for the purpose of egalitarian treatment. Logically, therefore, because there is no specific category that displays identifiable features that would justify differential treatment, there is no justification for exemptions from the law. (Laborde 2017: 44)

Laborde rejects Dworkin’s view because she believes it suffers from “hidden tensions” and “internal contradictions” (Laborde 2017: 46; cf. Laborde 2014a). Dworkin maintains that religion is not uniquely special, but just one possible conception of the good; as such, Dworkin believes religion is not ethically salient at all, that is, religion does not deserve special ethical consideration. But Laborde shows that sometimes Dworkin does acknowledge that religion is ethically salient, when he suggests that in some cases it is wrongful for the state to indirectly burden religious individuals’ “sacred duties” (Laborde 2017: 47). But if this is the case, then religious practice *is* ethically salient. Dworkin, however, does not offer the grounds for distinguishing between important practices motivated by sacred duties and unimportant practices (Laborde 2017: 48–9).

The other liberal egalitarian strategy that Laborde rejects is the one that she labels the “mainstreaming strategy” (Laborde 2017: 50–61), based on the view of Christopher Eisgruber and Lawrence Sager (2007). Eisgruber and Sager hold that religious believers should enjoy equal, but not special freedom. Religion is special, but not uniquely so; concerns of free exercise of religion are on a par with equally important and protection-worthy interests, such as familial commitments or the needs of the disabled (Laborde 2017: 52). Thus, Eisgruber and Sager believe that providing religious exemptions is justified by a concern for antidiscrimination.

But which feature of religion is ethically salient for egalitarian and antidiscrimination purposes? One feature that Eisgruber and Sager highlight is the vulnerability of religious minorities to hostile and unjust treatment of the majority. Their main example of majority neglect is the *Employment Division v. Smith* case, in which two Native Americans were denied unemployment benefits because they were fired for ingesting peyote, a hallucinogenic drug used for Native American rituals.² But because a similar exemption was provided to a religious ma-

² *Employment Division, Department of Human Resources of Oregon v. Smith*, 494 U.S. 872 (1990).

majority group, namely, the exemption for ceremonial wine at Christian masses in Oregon, not exempting Native Americans is discriminatory in their view (Eisgruber and Sager 2007: 92; Laborde 2017: 54). In addition, Eisgruber and Sager try to explain why religion is morally salient for exemption: religious commitments are “deep”, similarly to other serious protection-worthy commitments like familial commitments and disability-related needs (Laborde 2017: 55).

While sympathetic to Eisgruber and Sager’s emphasis on differential treatment for religious majorities and religious minorities, Laborde rejects their view because she does not agree with analogizing religion with disability. Religious practices are fundamentally different, Laborde holds, because disability is by and large a negative state of the person that she wants to get rid of, whereas a religious practice is not considered a detriment that calls for compensation (Laborde 2017: 56).³

The third mainstream approach to religious exemptions is the view of Jocelyn Maclure and Charles Taylor (2011), which Laborde labels the “narrowing strategy” (Laborde 2017: 61). In Laborde’s view, Maclure and Taylor reduce the ethical salience of religion to conscientious duties (Laborde 2017: 61). For instance, they analogize religion with other non-religious conscientious commitments, like pacifists’ rejection of military service (Maclure and Taylor 2011: 89–90). Laborde, however, dismisses their approach because she thinks it is *too* narrow: conscience as the ground for providing religious exemptions is not satisfactory because many religious claims for exemption are much more practice-oriented than the ‘narrowing strategy’ suggests. She mentions devout Muslims and Catholics as relevant examples: if a devout Muslim observes Ramadan, or a devout Christian observes Lent properly, then these practices are “fundamentally about exhibiting the virtues of the good believer, living in community with others, and shaping one’s daily life in accordance with the rituals of the faith” (Laborde 2017: 66–7), rather than conscientious duties, strictly speaking.

As I mentioned earlier, Laborde’s alternative view to these mainstream approaches is the disaggregation strategy. Armed with this method, Laborde aims to keep what is best in earlier theories without inheriting their shortcomings. Laborde dismisses Dworkin’s strategy, for it does not find religion ethically salient. But while she finds Maclure and Taylor’s approach too narrow, she builds on their theory by adopting their two-pronged theoretical structure. According to this approach, what must be established in the first step is the ethical salience of religion, after which comes the question whether the given practice should be accommodated or not.

What makes religious practices ethically salient for Laborde vis-à-vis Eisgruber and Sager and Maclure and Taylor? Despite her denial of the conscientious duty approach of Maclure and Taylor, Laborde believes the strongest form of ethical salience is in the neighborhood of

³ In subsection 2.3, I examine this point in greater detail.

conscientious duties, as she thinks that what makes religion ethically salient is *integrity*. Religious practices, according to Laborde, are related to “integrity protecting commitments”, or IPCs (Laborde 2017: 197–217). Laborde identifies two types of IPCs: obligation-IPCs, which have a scope wider than Protestant conscientious duties (because they refer to embodied practices on top of beliefs) and identity-IPCs, which may concern any religious practice connected to the believer’s deeper sense of the self (Laborde 2017: 197–217). Either of these two grounds are sufficient for Laborde to give a pro tanto reason for accommodation.

Once the question of ethical salience is answered, the second step is to discover principles that inform us when IPCs should be accommodated. Laborde proposes two principles for justifying exemptions: *disproportionate burden* and *majority bias* (Laborde 2017: 217–38). The disproportionate burden principle, which aims to ensure that laws will “avoid disproportionately burdening certain kinds of commitments” (Laborde 2017: 221), has four elements: *directness*, *severity*, *aim of the law*, and *cost-shifting*. *Directness* “is measured in relation to the costs incurred by individuals in avoiding subjection to the law or regulation in the first place” (Laborde 2017: 221). The more direct a burden that a law places on the given religious practice, the more reason we have for exempting it from the requirements of that law. How *severely* a religious practice is burdened on Laborde’s view depends on how deeply the practice is rooted in the believer’s *perceived* (i.e. subjective) obligation (Laborde 2017: 223). The *aim of the law* criterion observes the extent to which a given law is central “in promoting egalitarian justice” (Laborde 2017: 225). Thus, “[t]he more tightly a law promotes a goal of egalitarian justice, and the more it requires universal and uniform compliance for its effectiveness, the less it will tolerate exemptions” (Laborde 2017: 225). Finally, *cost-shifting* aims to guarantee that exemptions will not cause unjust burdens to those not exempted by the law, but this must be balanced against the previous three considerations. Laborde concludes that whether “exemptions are compatible with justice” depends on “if the balance of these four reasons renders the burden disproportionate” (Laborde 2017: 228).

Laborde’s second principle for religious exemption is majority bias. This principle justifies exemptions where the background conditions favor a historically dominant religion (Laborde 2017: 229). This is a comparatively egalitarian principle that is clearly influenced by Eisgruber and Sager. Thus, for example, the preference of Muslims to leave their workplaces earlier on Friday can be justified on the grounds that the structure of the working week displays a majority (Christian) bias.

In the remainder of the paper, I will show that despite its elegance, Laborde’s theory faces serious challenges.

3. *The problem with Laborde's conception of the ethical salience of religion*

In this section, I will highlight three different problems with Laborde's approach. First (subsection 2.1), I will highlight some methodological disagreements I have with the two-pronged structure of Laborde's theory. Second (subsection 2.2), I show her fundamentally individualistic formulation of ethical salience, i.e. that the ethical salience of religion depends solely on some personal features of the religious believer, to be too narrow. Finally (subsection 2.3), I show her rejection of the analogy between religion and disability to be unconvincing; I demonstrate this by looking at my argument from earlier work that establishes this analogy.

3.1 *Some methodological disagreements*

As we have seen above, Laborde highlights two features of religious practices that make them ethically salient: obligation-IPCs and identity-IPCs (while giving stronger moral weight to the former). Thus, Laborde's formulation of ethical salience does not cover all religious practices. She bites the bullet here. Reflecting on the view of Andrew Koppelman, who considers all religious activities as potentially protection-worthy even if they are only motivated by "habit, adherence to custom or happy religious enthusiasm," Laborde asserts that practices which spring from these motivations are not eligible for accommodation (Laborde 2017: 216; cf. Koppelman 2006). Is the omission of non-integrity-based religious practices a problem? I believe it is. This stems from my methodological disagreements with Laborde.

First, I find the two-pronged structure of Laborde's theory too narrow. I understand why she and others like Maclure and Taylor believe that, for accommodation, we need to explain why a given religious preference is special compared to a non-religious one: if Sikhs' motivation for wearing a turban on motorcycles, which excludes the possibility of wearing a crash helmet, merely comes from happy religious enthusiasm, then it is hard to argue against Brian Barry's former colleague, who assured him that "nothing matches riding a Harley-Davidson at full throttle down a deserted freeway, and [...] a bare head is essential to the value of the experience" (Barry 2001: 47). Establishing ethical salience in the first step protects one's theory of religious accommodation from such counterarguments. But it also excludes the possibility that there can be reasons for accommodating a religious practice from happy religious enthusiasm that do not apply to people who prefer to ride a Harley down a deserted freeway bareheaded just for the sake of the special experience. In the next two sections, I show that, in some cases, we can find compelling reasons to accommodate religious practices without establishing the ethical salience of the practices first.

These compelling reasons, as I demonstrate in the next two sections, are sometimes related to religious individuals as members of re-

ligious groups, and other times to how the environment interacts with the given religious preference. But this analysis that lays emphasis on the person-environment interaction can only be done if our theory is highly sensitive to context. What is surprising is that the disaggregation approach promises exactly such deep context-sensitivity, but its Labordian formulation fails to fully deliver. Instead of the rich descriptions of empirical cases and their history, Laborde's approach simply remains a close cousin to Maclure and Taylor's conscience-based theory. Hence, my second methodological criticism is that Laborde fails to take advantage of the full potential of the disaggregation approach.

Let me illuminate this methodological criticism. I agree with Emanuela Ceva that the main novelty of Laborde's approach to religion is a methodological one, namely, her disaggregation approach, which aims to both identify the important values at play in the multiple contexts where religion is present and "unpack the religious phenomenon itself" (Ceva 2018: 819). This "unpacking" of religion as a method to me suggests, to use Jonathan Wolff's phrase, a "problem-driven" (bottom-up) approach rather than a "theory-driven" (top-down) approach (Wolff 2011). Wolff holds that with public policies, we should apply the problem-driven approach, which implies that "the first task is to try to understand enough about the policy area to be able to comprehend why it generates moral difficulties, and then to connect those difficulties or dilemmas with patterns of philosophical reasoning and reflection" (Wolff 2011: 9). The problem-driven approach fits nicely with Laborde's disaggregation approach—disaggregating religion as a phenomenon should be handled, I believe, by first examining the kind of problem we are facing regarding a given religious practice, and understanding all the complexities of the situation. By contrast, Laborde's theory of exemptions seems top-down: she makes great efforts in the book to provide a theory that integrates all occurring and possibly emerging religious accommodation cases. That is, if we have an accommodation claim, we should try to think about how it would go through the filters of ethical salience (IPCs) and Laborde's two principles of accommodation.

But given how complex and diverse the problem of religious accommodation is, I believe these integrative, top-down theories that aim to explain and justify all accommodation claims are doomed to fail. Instead, theorists should aim to provide partial, bottom-up explanations, or insights that focus on the problems on the ground. In other words, instead of trying to "paint the full picture", theorists of religious accommodation should aim to provide partial theories without which the full picture of religious accommodation would be incomplete.⁴

⁴ Here I refer to Calabresi and Malamed's famous work (Calabresi and Malamed 1972). Calabresi and Malamed aim to provide a partial legal theory of property rules and torts, without offering a comprehensive theory of tort law. As they put it, "[a]s Professor Harry Wellington is fond of saying about many discussions of law, this article is meant to be only one of Monet's paintings of the Cathedral at Rouen. To understand the Cathedral, one must see all of them" (Calabresi and Malamed 1972: 1089n2).

3.2 *Laborde's ethical salience is too individualistic*

Laborde's approach runs into further problems, which stem from one of the fundamental features of her formulation of ethical salience: it is primarily individualistic. Laborde assumes that if there is something ethically salient about religious practices, it must be located at the level of a person. So, if we want to learn about the ethical salience of a practice, we must direct our focus onto the relationship between the person and her commitments. But this individualistic focus cannot *always* account for why we want to accommodate religious practices. Let me highlight three aspects of religious accommodation which render an individualistic approach unconvincing. First, consider Peter Jones' point about 'innocent' reasons for the special legal protection of religion:

Religions are more likely than non-religious beliefs to throw up norms of conduct that clash with prevailing socio-economic arrangements, such as holy days, religious festivals, dress codes, sacred symbols, prayer times, dietary requirements, and so on. It is difficult to imagine a non-religious system of belief generating a similar range of norms, unless it was itself quasi-religious. (Jones 2012: 1)

Jones' observation highlights an important point, namely that religions can frequently be *comprehensive* in the sense that they regulate similar aspects of their followers' lives as the state does for its citizens. I hold that the presence of this kind of comprehensiveness can give compelling reasons for considering the accommodation of any religious practice that clashes with similar majority practices, without relying on the question of ethical salience. I think this non-individualistic perspective should have been on Laborde's radar. In other words, her analysis should be sensitive to whether the burdens falling on the exercise of a given religious practice are *intrinsic* to that practice, or they are *positional*, in which case the burden is due to an environment where the practice clashes with majority arrangements.

Second, a further difficulty is related to group-based religious and cultural accommodation. Laborde fails to recognize that *numbers can matter*, in terms of how many individuals seek accommodation. What certainly does not help Laborde here is her focus on providing exemptions, which in many cases does not involve significant cost-shifting. Of course, I do not want to suggest that she fails to consider the distributive aspects of religious accommodation entirely, since the last of the four components of her principle of disproportionate burden concerns cost-shifting. I also acknowledge that exempting certain practices from generally applicable laws is a problem for liberals.

But Laborde's focus is somewhat narrow in that it detracts attention from the fact that accommodation often requires *assistance* to claimants, which can entail the modification of the social and material environment. Frequently, for such reconstructive claims, idiosyncrasy is not enough to justify accommodation however ethically salient the claim might be. Of course, an immediate response from Laborde can be that the first prong concerns only *pro tanto* reasons for accommoda-

tion; whether the given (ethically salient) religious practice should be accommodated is a further question. An ethically salient idiosyncratic claim might not end up being eligible for exemption based on considerations of distributive justice. But then again, the accommodation of any religious practice, even those that Laborde takes not to be ethically salient, might be justified once there are enough people within a socially salient group who share the given practice.

Take Akiva Nof's case. During the Gulf War in 1990, the state of Israel provided gas masks to its citizens in order to protect them from possible gas attacks. Interestingly, the state manufactured special gas masks (which were twice as expensive) for its religious citizens who were bearded, since normal gas masks were useless for them (gas masks need to be airtight). But, Akiva Nof, a non-religious bearded citizen, also wanted a special gas mask for free, claiming that his beard was an essential part of his identity. He was denied, so his case ended up in Israeli Supreme Court, where he won.⁵ For the sake of the present argument, let's assume that Nof's motivation counts as an integrity protecting commitment (IPC), but the Orthodox Jews' preference for beards does not (they merely grow beards for happy religious enthusiasm or adherence to synagogue fashion). Thus, a Labordian theory would only focus on the question whether Nof's ethically salient practice is worthy of accommodation, or not.

The difficulty with Laborde's individualistic integrity view in such a scenario is that it is counterintuitive. For if someone's integrity is what makes a difference, then the special gas mask should not be manufactured in the first place because the Orthodox Jews' preference for beards is not ethically salient. Moreover, special masks would arguably not be manufactured for the sake of a single person,⁶ like Nof, however ethically salient his practice is.

In my view, it would be very counterintuitive to argue that designing and manufacturing a single special size mask, even if not very costly, is a requirement of justice. I believe it makes much more sense in this context to hold that once a sufficiently large group (having Labordian IPCs or not) carves out an accommodation, other persons who have a justifiable claim (possibly including conscientious objectors) can join that accommodation. But an individual, ethically salient practice or preference alone does not have this accommodation prompting force. That is, the ethical salience of the individual identity-IPC in this case does not trigger accommodation, unlike moral considerations from the group level. Or, if we want to stay loyal to the notion of ethical salience, in cases like those of the special gas masks, we can say that ethical salience of religion is generated by its group, not its individual dimension.

At this point, Laborde could object that if numbers matter in such a way, then we prove too much. The argument would suggest that any group, even with *prima facie* trivial characteristics, might be eligible

⁵ I take this example from Perez (2009).

⁶ Here I assume the problem of the economies of scale.

for accommodation, and it would not be obvious what is special about religion; that is, we end up with the unappealing position of dissolution. To reply to this worry, on the one hand, I do not find it counterintuitive that even the most mundane characteristic/practice can be accommodated if it is shared by enough people. Consider cyclists' preference for more bike lanes. This preference is triggered by the fact that bikers in a neighborhood have reached "critical mass". It would be unfair to deny this accommodation to them (while simultaneously provide it to people using alternative means of transport, e.g. car drivers).

On the other hand, this is not dissolution, at least not in the Dworkinian sense. Dworkin builds on the same two-level assumption that characterizes Laborde's theory: the first step establishes the ethical salience of a religious practice, as a necessary but not sufficient (*pro tanto*) reason for accommodation. If a given practice is not ethically salient then accommodation is not warranted; since the ethical salience of religion for Dworkin cannot be established as part of this first step, accommodation is unjustified. But my view is different. The ethical salience of a given practice is not necessary for justifying the accommodation—the reason for accommodation might stem from a different source.⁷ I also do not share Dworkin's dissolution approach because I do consider certain religious practices ethically salient (including Labordeian IPCs), but I simply hold that building our theory of accommodation on ethically salient individual practices would be too narrow. In addition, I think that comprehensiveness and sufficient numbers make religious groups special beyond the individual level of beliefs and practices and this can arguably justify accommodation.

Moreover, a third group-based reason for accommodation is *social salience*. According to Kasper Lippert-Rasmussen, "a group is socially salient if perceived membership of it is important to the structure of social interactions across a wide range of social contexts" (Lippert-Rasmussen 2013: 30). Unlike Laborde, I hold that this aspect of religious groups is important in the accommodations/exemptions context, since minority group membership and related religious practices can be subject to the majority's discriminatory treatment, similarly to racial groups or the disabled. In my view, non-accommodation can sometimes expressively harm religious minorities, by sending a message to claimants that their practice is less valuable than that of the majority merely on the grounds of their group membership. Thus, accommodation might be justified if it helps to prevent expressive harm to religious minorities. Laborde rejects this expressive-discrimination perspective in the context of exemptions/accommodations:

Consider [...] freedom of speech. In most cases, religious convictions should be analogized to ideas and opinions open to public critique. But when the target of the speech is not the beliefs themselves but some prejudiced, libel-

⁷ This can be the *mismatch* between individual characteristics and environmental requirements as I emphasize in Section 2.3, where I draw the disability analogy.

ous, or offensive characteristic attributed to the group itself, we can talk of hate speech via the racialization of religion, and the second dimension of religion comes to the fore. In cases of exemptions, however, I have suggested that the salient dimension of religion is neither “belief” nor “race.” It is, rather, what I have called integrity-protecting commitments. They are weightier than mere belief (which may or may not be connected to integrity). And they are not an externally assigned, racelike identity: they are commitments that individuals positively identify with, which should not be construed as a disability or disadvantage. (Laborde 2017: 224)

But as I emphasize elsewhere, neglecting a socially salient group’s crucial characteristics in designing public arrangements amounts to negligence and is discriminatory (Zala 2018: 819). The wrong-making feature of this kind of negligence is exactly that it causes an expressive harm to these groups, as in the case of racist policies. In the Sikh crash helmet case, as I demonstrate in my other work, this element of expressive harm is certainly present, arguably justifying the Sikhs’ exemption from wearing crash helmets on motorcycles (Zala 2017: 11), whereas an individualistic, integrity-based notion of accommodation is counterintuitive in this case.

One might think that the role of Laborde’s principle of majority bias is exactly to tackle these problems. After all, this principle deals with individuals as members of religious groups. But it is puzzling why Laborde does not entertain the idea of expressive harms if she thinks majority bias is a crucial problem regarding exemptions. Ideally, the principle of majority bias should counter tendencies of indirect religious discrimination, which should not only showcase the relevance of expressive harms, but also highlight why establishing ethical salience is unnecessary when minority religious practices are proxies for treating groups unfairly. It also points to a fundamental tension in Laborde’s theory between the two prongs simply because any, even the most trivial religious practice, can be subject to majority bias. But if that is the case, then we do not have to treat ethical salience as a lexically prior entry requirement for accommodation. In the following section, I show that even the majority bias principle’s scope is too narrow because—as we will see regarding human variation issues—the need for minority accommodation does not always entail bias on the part of the majority.

3.3 *Religion and disability*

My final critical remarks refer to Laborde’s dismissal of the analogy between religion and disability. Laborde’s discussion of this issue is sporadic and exclusively focuses on Eisgruber and Sager’s argument, although she is definitely not alone in rejecting this analogy (Barry 2001: 36–7; Dworkin 2002: 291–6). Laborde’s view regarding disabilities and the analogy between disability and religion is unsatisfactory in two respects.

First, her rejection of the analogy comes from a false understanding of why disability is a social problem. According to Laborde, “[i]t makes

a moral difference that people (by and large) positively endorse and embrace their religious convictions, whereas they (by and large) prefer not to suffer from a disability” (2017: 56). This characterization of disability is often called the ‘personal tragedy view’ and underlies the ‘medical model’ of disability (Barnes 2016: 168). But the personal tragedy view is both offensive to people with physical impairments and misleadingly essentializes their situations. It is true that, for many people, losing an already possessed physical function is tragic and a circumstance of their lives that they would like to prevent from ever happening if they had the opportunity. This is especially true for those individuals whose life projects crucially depend on some physical characteristics and skills. Just think of a conductor who loses his hand in an accident (provided s/he does not possess the unique skills of Leonard Bernstein).

But this is not necessarily true for every person who is physically impaired. For those who were born with a functional limitation, this might not be a problem (that they should “overcome”) at all. This is because for many physically impaired people, their impairments are not, to use Ronald Dworkin’s distinction, *limitations* on but *parameters* for their lives and life plans (Dworkin 2002: 260). Laborde seems to acknowledge this when she says that

One may object [...] that people may also embrace, and identify with, their state of pregnancy, age, or disability. This is correct, but it misses the point: there are separate reasons for accommodating such states of being or cycles of life. IPC exemptions, by contrast, arise principally out of a conflict between the law and a given belief or project. An incidental burden on an IPC should be construed not as a disadvantage worthy of compensation but instead as one of the costs of (well-ordered) freedom. (Laborde 2017: 219)

Two things must be said regarding Laborde’s rejoinder. First, here again, she affirms that what she has in mind is the medical model of disability, which holds that disability is a personal deficit that must be corrected or compensated. A criticism can be raised that her view on disability is *ableist*, claiming implicitly that being able bodied is a norm that physically impaired people should live up to. The first problem with her denial of the analogy between disability and religious accommodation is that it could only be true if ableism were correct.

But there is a further difficulty with Laborde’s view: because of her inappropriate understanding of disability, she does not acknowledge the *role of the environment* in creating impediments for both disabled and religious individuals. Laborde, just as Eisgruber and Sager, considers disability an individual characteristic. For Eisgruber and Sager, the analogy between disability and religion is intuitive, and they try to substantiate it by stating that for disabled individuals, their disability is a comparably “deep” and “serious” commitment or project (Eisgruber and Sager 2007: 101, 104; Laborde 2014b: 58). Laborde rightly rejects their reasoning because it is unnecessary for accommodating the disabled that their physical impairments should have a special meaning to them. This individual focus is misleading because many of the diffi-

culties the disabled face are results of unaccommodating environments.

In previous work, I show exactly how the environment creates impediments for both religious minorities and the disabled. There, I apply the human variation model of disability (HVM) to the problem of religious accommodation. According to the HVM, disability is a problem of mismatch between personal characteristics and the environment (Wasserman et al. 2016). According to the chief proponents of this model, Richard Scotch and Kay Schriener:

[D]isability could be defined as an extension of the variability in physical and mental attributes beyond the present—but not the potential—ability of social institutions to routinely respond[...] [T]he problems faced by people with disability might be seen as the consequence of the failure of social institutions (and their physical and cultural manifestations) that can be attributed to the institutions' having been constructed to deal with a narrower range of variation than is in fact present in any given population. (Scotch and Schriener 1997: 155; quoted in Zala 2018: 812)

In other words, the problem of the physically impaired is that they are an *atypical* minority, the social and physical arrangements of which are tailored to the characteristics of the majority. Thus, I approach the Sikh crash helmet case from a completely different angle than Laborde. The complication is not that Sikhs object to the general requirement of wearing a helmet, but that the British state failed to take into account in enacting the mandatory helmet regulation that there are citizens in the UK for whom wearing helmets on top of their turbans is not an option (cf. Poulter 1998: 292). This raises moral controversies because the negligence of the British state could potentially send the message that Sikhs are less important citizens than non-Sikhs.

Laborde could reply that her principle of majority bias can account for human variation cases. But this comes at a cost: to begin with, she has to acknowledge that Eisgruber and Sager's mainstreaming strategy often makes sense. Consequently, the principle of majority bias would have to be refined, as it should no longer be about whether there is fairness among the treatment of religious groups, but about fairness between mainstream society and a religious minority. Or, if we are sufficientarians or capabilityarians, we do not even need a comparison group to justify reconstruction. We would hold that Sikhs must have the opportunity to ride motorcycles safely, independently of the fact that this opportunity is provided to others.

Consider the following example, involving the Japanese restroom experience of Irving Zola, a disability scholar and activist, who upon entering a public restroom in Japan was surprised to find not only that the restroom had toilets for the disabled “with an expansiveness in size and features” he has “never encountered outside of a private home”, but that the restroom also had a full range of accommodating toilets for “normal” people (Zola 1993: 23). Thus, the restroom had both traditional Japanese toilets (where one squats over a “squat hole”) and Western style toilets (Zola 1993: 23). Zola was amazed because he real-

ized he had been trapped in the dichotomous view of disability (according to which someone is either disabled, or not). The amazing variation of different toilets from the Japanese experience made him recognize that disability is a gradual notion, a matter of degree (Zola 1993: 23–4).

Yet, the reason why this example is interesting to me is because it shows a number of important things about accommodation in general. It not only shows that modifying the environment can be required by reasons for accommodating the needs of religious, cultural and physically impaired people, but that it would also be offensive if the characteristics of one of the socially salient groups were not considered during the design of the public restrooms.

We can also test Laborde's IPCs with Zola's example. First, the example shows that the need for accommodation depends neither on obligation-IPC's nor on identity-IPC's. Accommodation is triggered, I believe, because a lack thereof would send the message towards salient social groups, like the Japanese or the physically impaired, that the state does not consider them likely users of such public facilities (cf. Anderson 2010: 92). The moral assessment of this situation would not change if an additional religious group appeared requiring a new design for the toilet. Their preference for the special design of toilets would provide pro tanto reasons for accommodation even if it was neither an obligation, nor a matter of personal identity, but merely an ordinary, trivial reason.

4. Conclusion

In this paper, I showed that Laborde's theory of religious exemptions faces some important difficulties. These are related to her formulation of the ethical salience of religious practices that provide reasons for accommodating such practices exclusively. I pointed out that her emphases on obligations and identity as the grounds of religious individuals' integrity, which, in turn, serves as the backdrop for accommodation, faces some challenges. First, her formulation of the ethical salience of religious practices based on personal integrity is too narrow. Second, her individualistic understanding of ethical salience faces several problems. Finally, I aimed to show that her rejection of the analogy between religious and disability accommodation is unconvincing.

References

- Anderson, E. 2010. "Justifying the capabilities approach to justice." In H. Brighouse and I. Robeyns (eds.). *Measuring Justice: Primary Goods and Capabilities*. Cambridge: Cambridge University Press: 81–100.
- Barnes, E. 2016. *The Minority Body*. Oxford: Oxford University Press.
- Barry, B. 2001. *Culture & Equality: An Egalitarian Critique of Multiculturalism*. Cambridge: Harvard University Press.
- Calabresi, G. and Melamed, A. D. 1972. "Property rules, liability rules,

- and inalienability: one view of the cathedral." *Harvard Law Review* 85: 1089–1128.
- Ceva, E. 2018. "Cécile Laborde, Liberalism's Religion." *Ethics* 128 (4): 819–823.
- Dworkin, R. 2002. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge: Harvard University Press.
- 2013. *Religion Without God*. Cambridge: Harvard University Press.
- Eisgruber, C. and Sager, L. 2007. *Religion and the Constitution*. Cambridge: Harvard University Press.
- Jones, P. 2012. *Discrimination, Exemption, and Non-Religious Belief*. Conference paper handout, Non-Religious Identities in Policy and Practice, 20 April 2012, Senate House, London. Available at: http://www.religionandsociety.org.uk/uploads/docs/2012_06/1339510501_Discrimination,_Exemption,_and_Non-Religious_Belief.pdf
- Koppelman, A. 2006. "Is it fair to give religion special treatment?" *University of Illinois Law Review* 3: 571–603.
- Laborde, C. 2014a. "Dworkin on freedom of religion without religion." *University of Boston Law Review* 94: 1255–71.
- 2014b. "Equal liberty, nonestablishment, and religious freedom." *Legal Theory* 20: 52–77.
- 2015. "Religion in the law: the disaggregation approach." *Law and Philosophy* 34: 581–600.
- 2017. *Liberalism's Religion*. Cambridge: Harvard University Press.
- Levy, J. 1997. "Classifying cultural rights." *NOMOS XXXIX: Ethnicity and Group Rights* 39 (9): 24–66.
- Lippert-Rasmussen, K. 2013. *Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination*. Oxford: Oxford University Press.
- Maclure, J. and Taylor, C. 2011. *Secularism and Freedom of Conscience*. Cambridge: Harvard University Press.
- Perez, N. 2009. "Cultural requests and cost internalization." *Social Theory and Practice* 35 (2): 201–228.
- Poulter, S. 1998. *Ethnicity, Law, and Human Rights*. Oxford: Oxford University Press.
- Scotch, R. K., and Schriner, K. 1997. "Disability as Human Variation: Implications for Policy." *The Annals of the American Academy of Political and Social Science* 549 (1): 148–159.
- Wasserman, D. et al. 2016. "Disability: Definitions, Models, Experience." *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2016/entries/disability/>.
- Wolff, J. 2011. *Ethics and Public Policy: A Philosophical Inquiry*. New York: Routledge.
- Zala, M. I. 2018. "The devout and the disabled: religious and cultural accommodation-as-human-variation." *Journal of Applied Philosophy* 35 (4): 809–824.
- Zola, I. K. 1993. "Disability Statistics, What We Count and What It Tells Us: A Personal and Political Analysis." *Journal of Disability Policy Studies* 4 (9): 9–39.

Procreative Beneficence toward Whom?

SERGIO FILIPPO MAGNI
University of Pavia, Pavia, Italy

This article deals with a discussion of Savulescu's impersonal version of the Principle of Procreative Beneficence and its relationship with a person-affecting Principle of Harm in order to evaluate the cases of selection of which child to have. It aims to show some problems in Savulescu's attempt to arrange the two principles (the conflict between beneficence and harm, the limitation of beneficence to pre-conception selection, the extension of beneficence to different quantity people choice), and to propose an alternative version of Procreative Beneficence (a narrow person-affecting version), in order to avoid these problems.

Keywords: Procreative beneficence, selection, impersonal obligations, person-affecting obligations, possible people.

1. *Beneficence and harm*

Julian Savulescu has stated an important moral principle related to reproductive decisions concerning cases of children's selection, at present related to post-conception testing, like chorionic villus sampling or amniocentesis, or to pre-conception testing, like Pre-implantation Genetic Diagnosis (PGD) and IVF.

This principle is a consequentialist principle, called the Principle of Procreative Beneficence (PB). During the years, it has had different formulations:

Couples (or single reproducers) should select the child, of the possible children they could have, who is expected to have the best life, or at least as good a life as the others, based on the relevant, available information. (Savulescu 2001: 415)

Or, according to a more recent formulation:

If couples (or single reproducers) have decided to have a child, and selection is possible, then they have a significant moral reason to select the child, of the possible children they could have, whose life can be expected, in light of the relevant available information, to go best or at least not worse than any of the others. (Savulescu and Kahane 2009: 274)

According to this principle, there is a moral obligation to have the most advantaged child (the best child) of all the possible children that a couple can have. PB is a consequentialist maximizing principle: it states an obligation to have the child with the most well-being (a life with Memory, Intelligence, Empathy, etc.); and it is a comparative principle: it compares the well-being of the possible children the parents could have (Savulescu and Kahane 2009: 175).

A first example of this principle's application regards a case of selection related to a pre-conception test for asthma:

A couple is having IVF in an attempt to have a child. It produces two embryos. A battery of tests for common diseases is performed. Embryo A has no abnormalities on the tests performed. Embryo B has no abnormalities on the tests performed except its genetic profile reveals it has a predisposition to developing asthma. Which embryo should be implanted? Embryo B has nothing to be said in its favour over A and something against it. Embryo A should (on pain of irrationality) be implanted. (Savulescu 2001: 416)

A second example regards a case of selection related to a pre-conception test during rubella (an example derived from Parfit):

A woman has rubella. If she conceives now, she will have a blind and deaf child. If she waits three months, she will conceive another different but healthy child. She should choose to wait until her rubella is passed. (Savulescu 2001: 417)

The Principle of Procreative Beneficence states an important view in reproduction ethics, which is opposed to several anti-selection views (like that of the Roman Catholic Church or of other secular thinkers, like Sandel): "far from playing God, attempting to control our genetic fate is 'playing human'—trying to improve the odds of doing well in an uncertain world of difficulty, threat and misfortune" (Savulescu 2007: 284).

But PB is also opposed to some pro-selection views (like those of Glover or Agar) which give the couple moral permission, but not a moral obligation, to select. These views *deny* that we ought to select the most advantaged child, because they believe that reproduction is a private matter; or because they think that morality gives parents complete freedom when they make procreative decisions, allowing them to aim at less than the best (Savulescu and Kahane 2009: 175).

Yet, according to Savulescu, there are limits on Procreative Beneficence: the obligation is not an absolute obligation but a *prima facie* (a *pro tanto*) obligation, which has to be balanced with other *prima facie* obligations in order to become an effective, "all-things-considered overriding obligation" (Savulescu and Kahane 2017: 594). For this reason, Savulescu states that, in PB, 'should' means 'have good reasons to' or 'is morally required', and it is different from 'must'. It implies that persuasion is justified but coercion is not.

PB is not an absolute obligation. It is the claim that there is a *significant moral reason* to choose the better child. The principle states not what people invariably must do but what they have significant moral reason to do. (Savulescu and Kahane 2009: 278)

In particular, it has to be balanced with another fundamental principle: the Principle of Harm, that states a moral obligation not to harm others. This requirement was originally related, in Savulescu's early articles, to a principle of Procreative Autonomy: "people should be free to do what others disapprove of or judge wrong, provided the exercise of freedom does not harm others" (Savulescu 2002). Explaining this principle, Savulescu refers directly to Mill's liberalism, according to which an invasion of a person's liberty is justified only when that person threatens to directly and seriously harm another person (Mill's Principle of Harm) (Savulescu and Kahane 2017: 612).

Following such a non-harming restriction, Savulescu admits the couple's free choice to select less than the most advantaged child. "There are strong philosophical grounds to hold that procreative liberty should extend to people selecting less than the best child" (Savulescu 2014: 178).

Not surprisingly, there may be a conflict between the two principles, and the solution proposed by Savulescu is to give precedence to the Principle of Harm, as we will see below. The result is "a liberal form of procreative perfectionism", as Glover (2006: 54) defined Savulescu's proposal, or a form of "libertarian consequentialism".

2. *A hybrid position*

Like many other decisions in reproductive ethics, selection changes the identity of the children who will exist: "in selecting a more advantaged child we are also bringing a different person into existence" (Savulescu 2009: 277).

In this way, selection is related to a well-known problem regarding future generation ethics: the *non-identity* problem, due to the fact that our actions can change the identity of future people. Therefore, selection brings with itself the need to distinguish different ways of understanding normative principles applied in reproduction ethics: a person-affecting view and an impersonal view.

On the one hand, in Savulescu, the Principle of Harm is intended as intrinsically related to a person-affecting restriction (in a narrow sense, as we will see later): "harm is personal in nature" (Savulescu 2014: 178).

A "Person-affecting View" states that an act is right or wrong only if there is or will be a person affected by it" (Glover 1977: 66). According to this view: we have an obligation only toward actual or future people who will exist. As Narveson (1967: 63) writes: "in deciding what we are to do, the only consideration which is morally relevant [...] is how others would be affected. If we cannot envisage effects on certain people which would ensue from our acts, then we have no moral material to work on and we can do as we like". Wrongs require victims: "what is bad must be bad for someone" (Parfit 1984: 363).

According to this person-affecting view, we do not benefit a child if we bring him/her into life, but if his/her being alive is not worth living we harm a child by procreating him. “If, therefore, it is our duty to prevent suffering and relieve it, it is also our duty not to bring children into the world if we know that they would suffer or that we would inflict suffering upon them” (Narveson 1967: 71).

On the other hand, in Savulescu, the Principle of Procreative Beneficence explicitly denies a person-affecting restriction: it is *impersonal*. According to such a view, an act is right or wrong even if there is or will be no person affected by it, but he/she could have existed if we had chosen otherwise. According to this view, we have an obligation even toward ‘merely possible people’, people who will not exist but whose existence depends on us; and a harm made to actual or possible people may be compensated for with a benefit made to merely possible people and vice-versa (Hare 1988: 68; Kamm 2013: 291).

In this impersonal interpretation of PB, we have a different kind of harm: an impersonal harm, a harm without a victim, which makes the world a worse place. As Savulescu states: “It is bad that blind and deaf children are born when sighted and hearing children could have been born in their place” (Savulescu 2001: 423).

Therefore, in Savulescu’s view the conflict between Procreative Beneficence and Harm becomes a conflict between an impersonal Principle of Procreative Beneficence and a person-affecting Principle of Harm. And Savulescu solves the conflict by stating the priority of a person-affecting harm on an impersonal beneficence:

In general, proscriptions against person-affecting harm should be much stronger than those against impersonal harms. We should not significantly harm individuals now to bring about a better world, where the betterness is impersonal. Impersonal reasons then could be rather weak when pitted against personal reasons. Requirements of PB [...] are reasons, but should not require major person-affecting harms. (Savulescu 2014: 178)

Therefore, Savulescu’s proposal regarding selection is a hybrid position, which admits both impersonal and personal moral reasons for actions (and therefore impersonal and personal moral obligations) and states an order of priority among them (for other hybrid positions in future generation ethics see Mulgan 2006: 82 ff): “In my view, there are both person-affecting reasons and impersonal reasons. However, person-affecting reasons are stronger than impersonal reasons, even though the latter are reasons for action” (2014: 177). But, when “there are no competing person affecting reasons, then PB should decide the day” (Savulescu and Kahane 2017: 612).

In order to clarify this position, Savulescu refers to Parfit’s example of the Two Medical Programmes (one for the treatment of a disease, the other for the prevention of a disease) in order to refuse Parfit’s solution and to state the priority of person-affecting reasons over impersonal reasons: “personal harms are much worse than impersonal ones” (Savulescu 2014: 177).

In this example, there are two medical programmes to detect two rare diseases, *J* and *K*, in women. If a pregnant woman has disease *J*, this will cause the child to have a certain handicap, but a simple treatment would prevent this effect. If a non-pregnant woman has disease *K* but decides to conceive a child, this will cause the child to have the same handicap; disease *K* cannot be treated, but it can be prevented, because it always disappears within two months. Unfortunately, there are funds for only one programme. Parfit supports the No Difference View between the two programmes: he believes that each programme is equally right because they have the same consequences in terms of handicapped people. On the contrary, according to Savulescu, there is a difference between the two programmes:

in the case of *Treatment*, there is person-affecting harm. If you fail to treat, a future person is made worse off than he or she would otherwise have been. In *Prevention*, the harm is impersonal—the world is worse for having more suffering than it could have contained, but no person is worse off than he or she would otherwise have been. (Savulescu 2014: 176–7)

This person-affecting limitation reduces the revolutionary impact of Procreative Beneficence and diminishes its radicality, permitting a general agreement with common-sense morality. As Savulescu underlines: “although PB is often presented as a radical view, it is really just an extension of widely accepted existing practices and an application of common-sense ethical ideas” (Savulescu and Kahane 2017: 598).

For this reason, as we have seen, deaf or dwarf people should be allowed to be free to deliberately select children with deafness or dwarfism:

What if a couple has in vitro fertilisation and preimplantation genetic diagnosis and they select a deaf embryo? Have they harmed that child? Is that child worse off than it would otherwise have been (that is, if they had selected a different embryo)? No—another (different) child would have existed. The deaf child is harmed by being selected to exist only if his or her life is so bad it is not worth living. Deafness is not that bad (Savulescu 2002: 772).

3. *Some problems*

Savulescu’s theory of selection has had different kinds of critics. Some scholars have refused Savulescu’s position from an anti-selection point of view, which does not consider selection of the best child as a morally permissible act (because of problems related to selection procedures, De Melo-Martín 2004 and 2016, or related to identifying what a best life is, Parker 2007). Others have refused Savulescu’s position from a liberal pro-selection point of view, which does not consider selection of the best child as a moral obligation but only as a moral permission, not correlated to any duty of beneficence (see Herisonne-Kelly 2006, Bennett 2008, Sparrow 2014). Others have refused Savulescu’s position from anti-consequentialist points of view (see Stoller 2008, Hotke 2012).

Here, I will focus on some problems related to the hybrid aspect of the theory. “If our reasons—Savulescu writes—to have the best chil-

dren (and best future environment) are relatively weak, they are easily overridden by person-affecting reasons” (2014: 178; Savulescu and Kahane 2017: 612). But, how rigid is this priority of a person-affecting harm over an impersonal beneficence? Does this hierarchical priority of person-affecting reasons over impersonal reasons need to be affirmed in every case of personal harm? How should we evaluate a selection in cases where the foetus shows severe handicap (but his life is however worth living)? In these cases, can beneficence have a stronger influence than harm when the severity of the handicap is relevantly high? Namely, how should we decide in borderline cases? There is a need for further specification of the theory and, probably, to set some limitations on the prevalence of the person-affecting principle of harm.

Secondly, it is better to distinguish the moral evaluation of the two different kinds of selection: in pregnancy (correlated to post-conception testing) and before pregnancy (correlated to pre-conception testing). Despite Savulescu’s comprehensive treatment of the moral evaluation of selection, Procreative Beneficence becomes an all-considered moral obligation only in the case of pre-conception testing (PDG), but not in the case of post-conception testing (chorionic villus sampling or amniocentesis), because of the priority given to person-affecting harm. Given that a foetus should not be harmed, as in Savulescu’s refusal of the No Difference View in the Two Medical Programmes example, selection against the foetus in a case of pregnancy causes a personal harm to the foetus (killing him/her) and therefore should be avoided. In such a case impersonal reasons are always weaker than personal reasons.

More in general, due to the priority given to the Principle of Harm, in every case of pregnancy the Principle of Procreative Beneficence does not result to be, *de facto*, effective, and it cannot be used as a means to address other reproductive issues, for example the moral question of abortion. A severe limitation for a principle that claims to be a guiding principle in reproductive ethics. On the contrary, following an impersonal maximizing consequentialist view in a case of abortion and an impersonal conception of harm, Hare has compared the quality of the probable life of the foetus with that of a possible future child, and on this basis has admitted some cases of abortion:

If the present foetus is going to be miserably handicapped if it grows into an adult, perhaps because the mother had rubella, but there is every reason to suppose that the next child will be completely normal and as happy as most people, there would be reason to abort this foetus and proceed to bring to birth the next child, in that that next child will be much gladder to be alive than will this one. (Hare 1993: 157–8)

Because of the prevalence of the person-affecting Principle of Harm, this pro-abortion argument based on an impersonal conception of harm is precluded to Savulescu.

Thirdly, as an impersonal maximization principle of well-being, PB should be extended to *different quantity* people choice, even in the case of selection. A maximizing conception of beneficence requires that less

quality should be compensated with more quantity: two less good is better than one good. So, in a case of Pre-implantation Genetic Diagnosis, impersonal maximizing Procreative Beneficence requires that we should implant the highest number of healthy embryos in order to have the highest number of future healthy children and the greatest quantity of well-being. Similarly, in a case of post-conception testing on two asthmatic twin embryos it requires that we should select in favour of the two asthmatic embryos rather than in favour of one future healthy child.

This extension from quality to quantity complicates the matter, because it implies other problems: the moral obligation to increase the number of healthy children, namely to procreate as many children as possible (up to Parfit's Repugnant Conclusion on the increasing population), and the necessity to distinguish a Total from an Average version of the Principle of Beneficence to try to avoid this problem. That is, it brings with it all the puzzles of future generation ethics (see also Bennett 2014).

Indeed, Savulescu explicitly limits the Principle of Procreative Beneficence to *same quantity* people choice: "it is a claim only about same number choice: about selection of one child out of those possible. It is not meant to offer guidance in choices between, e.g. one versus several children of differing endowments" (2009: 274). But this limitation seems to be an *ad hoc* manoeuvre, just in order to avoid these problems. If PB is a consequentialist maximizing principle, as Savulescu states, it is implausible to exclude *a priori* any extension to *different quantity* people choice, if the principle purports to remain a maximizing principle.

4. *Person-affecting procreative beneficence*

As proposed by Savulescu, the Principle of Procreative Beneficence is impersonal, but it could have different interpretations. Savulescu recognizes two different ways of interpreting it:

like competing principles of procreative ethics, PB is compatible with different accounts of reasons to select future children. It can take either a wide person-affecting form or an impersonal form. According to the wide person-affecting version, our reason to select the child with better prospects is that that child will benefit more than the other would by being caused to exist. According to the impersonal version, our reason is that selecting the most advantaged child would make the outcome better, even if it is not better for the child created. It is possible to support PB on either view (Savulescu and Kahane 2009: 277).

The two interpretations have the same consequences in terms of moral consideration of selection, and Savulescu opposes these versions of PB against the perspective contrary to selection (the Anti-Selection view) and other pro-selection perspectives (the Procreative Autonomy view, the Minimal Threshold view, the Satisficing view and the Prevention of Harm view) (Savulescu and Kahane 2009: 289).

But there is another possible view of selection (and more generally of reproductive ethics) which can avoid the problems seen above (the conflict between an impersonal beneficence and a personal harm, the limitation of beneficence to pre-conception selection, the extension of beneficence to different quantity people choice) without renouncing a maximization principle: a narrow person-affecting view of Procreative Beneficence, which aims to maximize the well-being of all the present or future *existing* persons affected by the action.

According to such a narrow person-affecting view (Parfit 1984: 395), in the case of selection Procreative Beneficence states that (following Savulescu's formulations):

Couples (or single reproducers) should *not select* the child, of the possible children they could have, who is expected to have a life which is *for him/her* not worth living (in terms of lack of well-being), based on the relevant, available information.

Or, in the other formulation:

If couples (or single reproducers) have decided to have a child, and selection is possible, then they have a significant moral reason not to select a child, of the possible children they could have, whose life can be expected, in light of the relevant available information, to be *for him/her* not worth living (in terms of lack of well-being).

Such a principle refers to both kinds of selection: selection in pregnancy (correlated to post-conception testing) and before pregnancy (correlated to pre-conception testing), because it exclusively considers the well-being of the child who will exist. Secondly, it avoids the conflict between impersonal beneficence and person-affecting harm because, in the case of selection, it does not need to make recourse to an obligation not to harm others in order to agree with common-sense morality. Thirdly, it avoids the problem of increasing the quantity of happy people, because for the person-affecting restriction there is no obligation to procreate the highest number of happy children.

The consequences of such a narrow person-affecting principle of Procreative Beneficence are, in part, different from Savulescu's view: there is no moral obligation for the couple to select the best of the possible children, but only moral permission. In the example of rubella, such a view does not say to the woman that she *should* wait three months, but that she *could* (not taking into consideration other prudential or moral reasons for waiting, which may be decisive).

In part, the consequences are similar to Savulescu's view: both allow several selective actions to be morally permitted (for example, to select in favour of an asthmatic child or a dwarf child), but not other selective actions (for example, to select in favour of an anencephalic child (Savulescu-Kahane 2017: 597)). According to this narrow person-affecting version of Procreative Beneficence, parents are free to select the best or someone less than the best, but they are not free to select

the worst, because such a version of PB remains a consequentialist maximizing view.

A narrow person-affecting Procreative Beneficence is a pro-selection view that is opposed to anti-selection views which do not permit selection, but it is less strong than an impersonal (or wide person-affecting) PB, because—some extreme cases apart—it does not consider selection as a moral obligation but only as moral permission.

Yet, it is stronger than other pro-selection views, such as Procreative Autonomy, because it does not allow a selection in favour of the worst and, when it is extended beyond selection and beyond reproductive choice, it transforms a negative duty (the obligation not to select) into a positive duty: the obligation to produce the best possible life for the child selected, maximizing his/her well-being (genetic enhancement included), because it remains a consequentialist maximizing principle of beneficence. According to such a personal PB, for example, we have an obligation to make a foetus immune to dangerous viruses through genetic intervention, when it is possible and safe.

Differently from Procreative Autonomy and like impersonal Procreative Beneficence, such a view “is not indifferent to the choice between a child who will have a fulfilling life and one who will live a brief life of misery and torment” (Savulescu and Khane 2009: 279).

References

- Bennett, R. 2008. “The fallacy of the principle of procreative beneficence.” *Bioethics* 23 (5): 265–273.
- Bennett, R. 2014. “When intuition is not enough. Why the principle of procreative beneficence must work much harder to justify its eugenic vision.” *Bioethics* 28 (9): 447–455.
- De Melo-Martín, I. 2004. “On our obligation to select the best children: a reply to Savulescu.” *Bioethics* 18 (1): 72–83.
- De Melo-Martín, I. 2016. *Rethinking Reprogenetics: Enhancing Ethical Analyses of Reproductive Technologies*. Oxford: Oxford University Press.
- Glover, J. 1977. *Causing Death and Saving Lives*. London: Penguin.
- Glover, J. 2006. *Choosing Children. Genes, Disability and Design*. Oxford: Oxford University Press.
- Hare, R. M. 1975. “Abortion and the Golden Rule.” In Hare 1993: 147–167.
- Hare, R. M. 1988. “Possible People.” In Hare 1993: 67–83.
- Hare, R. M. 1993. *Essays on Bioethics*. Oxford: Oxford University Press.
- Herisonne-Kelly, P. 2006. “Procreative Beneficence and the Prospective Parents.” *Journal of Medical Ethics* 32 (3): 166–169.
- Hotke, A. 2012. “The Principle of Procreative Beneficence: Old Arguments and a New Challenge.” *Bioethics* 28 (5): 255–262.
- Kamm, F. 2013. *Bioethical Prescriptions to Create, End, Choose, and Improve Lives*. Oxford: Oxford University Press.
- Mulgan, T. 2006. *Future People. A Moderate Consequentialist Account of our Obligations to Future Generations*. Oxford: Clarendon.

- Narveson, J. 1967. "Utilitarianism and New Generations." *Mind* 76 (30): 62–72.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon.
- Parker, M. 2007. "The best possible child." *Journal of Medical Ethics* 33 (5): 279–283.
- Savulescu, J. 2001. "Procreative beneficence: Why we should select the best children." *Bioethics* 15 (5–6): 413–426.
- Savulescu, J. 2002. "Deaf lesbians, 'designer disability' and the future of medicine." *British Medical Journal* 325 (7367): 771–773.
- Savulescu, J. 2007. "In defence of procreative beneficence." *Journal of Medical Ethics* 33 (5): 284–288.
- Savulescu, J. 2014. "The nature of the moral obligation to select the best children." In A. Akabayashi (ed.). *The Future of Bioethics. International Dialogues*. Oxford: Oxford University Press: 170–182.
- Savulescu, J. and Kahane, G. 2009. "The moral obligations to create children with the best chance of the best life." *Bioethics* 23 (5): 274–290.
- Savulescu, J. and Kahane, G. 2017. "Understanding procreative beneficence." In L. Francis (ed.). *The Oxford Handbook of Reproductive Ethics*. Oxford: Oxford University Press.
- Sparrow, R. 2014. "Ethics, Eugenics and Politics." A. Akabayashi (ed.). *The Future of Bioethics. International Dialogues*. Oxford: Oxford University Press: 152–178.
- Stoller, S. E. 2008. "Why we are not morally required to select the best children: A response to Savulescu." *Bioethics* 22 (7): 364–369.

Michael Walzer's Republican Theory of Distributive Justice: "Complex Equality" as Equal Freedom from Domination

MANUEL KNOLL

Istanbul Şehir University, Istanbul, Turkey

This article presents a republican interpretation of Michael Walzer's theory of distributive justice and of his idea of complex equality. It demonstrates that Spheres of Justice is not only a defense of pluralism and equality (as the subtitle announces), but also of liberty or freedom. Like Quentin Skinner and Philip Pettit, Walzer understands liberty as non-domination. For Walzer, a just distribution of all social goods leads to a "complex egalitarian society" in which every citizen is equally free from domination and tyranny. Against alternative interpretations, this paper suggests that Walzer is indeed a political egalitarian and that complex equality should be interpreted as a simple equality of liberty or freedom. In the conclusion, the article argues that Walzer's and Pettit's versions of republicanism are complementary because they each illuminate the other's blind spot and thus mutually fix each other's particular shortcoming.

Keywords: Republicanism, liberty, non-domination, equality, egalitarianism, tyranny.

1. Rawls and Walzer

Almost 50 years have passed since John Rawls's *A Theory of Justice* was published in 1971.¹ This influential work, in which Rawls tries to substantiate two principles of justice for social institutions, has led to a true renaissance of normative political philosophy. Still today, his theory "dominates contemporary debates, not because everyone accepts

¹ In the first months of 1975 Rawls added some changes and improvements in his work for the German Edition which came out in 1979. This version was not published in English until 1999.

it, but because alternative views are often presented as responses to it" (Kymlicka 2002: 10). In some ways this is also true for Michael Walzer's groundbreaking work *Spheres of Justice*, which came out in 1983. Though Walzer contends that his "enterprise is very different from Rawls's," he admits that it "would not have taken shape as it did—it might not have taken shape at all—without his work" (Walzer 1983: xviii).

In the preface of the German edition of *Spheres*, Walzer explains what he holds to be the main difference between his and Rawls's theory. According to Rawls, his two principles of justice are sufficient to regulate the distribution of all desirable social goods, like liberty, opportunity, income and offices. Against this claim, Walzer argues that the broad range of different social goods—membership, welfare, security, free time, education, recognition, political power, etc.—cannot be reduced to "a short list of basic goods," and neither are two principles of justice sufficient to regulate the just distribution of all these social goods (Walzer 1983: 4; Walzer 1992: 12). Rather, Walzer calls for a diverse set of rules, standards and principles for the distribution of all different social goods. While "from Plato onward," the majority of philosophers who have written about justice assume that "there is one, and only one, distributive system," Walzer argues for a pluralist approach that encompasses a variety of distributions and distributive principles. He claims "that the principles of justice are themselves pluralist in form; that different social goods ought to be distributed for different reasons, in accordance with different procedures, by different agents" (Walzer 1983: 5–6). For him, there is only one universal procedural rule: each social good should be distributed according to the criteria valid for its own sphere (Walzer 1992: 12).

Walzer's main suggestion for the multiplicity of social goods and the complexity of distributive systems is his idea of "complex equality". This remarkable idea reconciles the common egalitarian demand for social equality with the recognition of a large number of social inequalities. Walzer distinguishes the political egalitarianism he advocates from the prevailing egalitarian approaches that aim at establishing "simple equality" (Walzer 1983: 13–17). Simple egalitarians usually focus on one social good, like resources or welfare, and argue that this good should be redistributed towards the goal that everyone has the same amount of this good.² Contrary to this, Walzer's claim for "complex equality" permits unequal distributions of social and economic

² In his first two articles on equality Ronald Dworkin focuses on the problem of "distributional equality", which does not concern the distribution of political power but of money and other resources to individuals (Dworkin 1981a, 1981b): "I shall consider two general theories of distributional equality. The first (which I shall call equality of welfare) holds that a distributional scheme treats people as equals when it distributes or transfers resources among them until no further transfer would leave them more equal in welfare. The second (equality of resources) holds that it treats them as equals when it distributes or transfers so that no further transfer would leave their shares of the total resources more equal" (Dworkin 2001: 12).

goods if these goods are distributed under certain conditions. This is also true for Rawls's difference principle which allows social and economic inequalities if they are to the greatest benefit of the least advantaged members of society.³

In a just society that Walzer conceptualizes as a "complex egalitarian society" (Walzer 1983: 17, 320), most goods are distributed unequally. This raises the questions why Walzer sees himself as a political egalitarian at all and why he calls such a society an "egalitarian society." What exactly does his idea of "complex equality" mean and what kind of equality does the term refer to? In the literature on Walzer, some interpretations of "complex equality" have been offered (Arneson 1995, Haus 2000, Miller 1995b, Swift 1995). Though grasping some important aspects of Walzer's idea, however, these approaches seem to miss his main point. The proposal here is for a different reading of Walzer's "theory of complex equality" (Walzer 1983: 28), an interpretation that might be surprising. As its main thesis, this paper suggests that complex equality should be interpreted as an equality of freedom or liberty. Contrary to John Rawls's conception of justice as fairness that demands primarily equal liberty, the idea and value of liberty seems to be of no significant importance in Michael Walzer's normative theory of justice. In *Spheres*, the terms "liberty" and "freedom" are mentioned only a few times, and the subtitle announces merely *A Defense of Pluralism and Equality*. Despite these facts, this paper aims to show that liberty and freedom play a key role in Walzer's theory of justice. Like proponents of contemporary republicanism, Walzer understands liberty primarily in terms of a state that protects all citizens in a reliable way from domination. Such an understanding, which defines liberty as non-domination, was elaborated by Quentin Skinner, and in particular by Philip Pettit in his two books *Republicanism: A Theory of Freedom and Government* (1997), and *On the People's Terms. A Republican Theory and Model of Democracy* (2012). A detailed consideration of Walzer's "complex equality" from the perspective of Pettit's republicanism leads to a better understanding of this idea.

Walzer himself makes clear that he is in favor of a republicanism that is adapted to the pluralism of civil society. Such a "pluralist republicanism," he suggests, "is also likely to advance the prospects of what I called 'complex equality'" (Walzer 2005b: 178, cf. 160–161; cf. Walzer 2007b: 116–120). However, usually Walzer refers to himself as a democratic socialist (Walzer 1980).⁴ In the literature, he is most commonly described as a communitarian (Avineri and de-Shalit 1992: 7,

³ The final statement of the difference principle in *A Theory of Justice* reads: "Social and economic inequalities are to be arranged so that they are [...] (a) to the greatest benefit of the least advantaged, consistent with the just savings principle" (Rawls 1971: 302).

⁴ In *Spheres*, Walzer states that the appropriate "institutional arrangement" for complex equality in "our own society" is a "decentralized democratic socialism" (Walzer 1983: 318).

10; Benbaji and Sussman 2014: 2; Kymlicka 2002: 209; Reiner 2011. Yet for his book *Politics and Passion* Walzer chose the subtitle *Toward a more Egalitarian Liberalism* (Walzer 2005a).⁵ In line with this subtitle, David Miller declares in the introduction of a book he edited with Walzer's articles: "Politically, Walzer is a liberal" (Walzer 2007a: xi–xii). There are good reasons for the claim that all these labels fit Walzer to some extent. Nevertheless, his idea of complex equality is most adequately understood through the lens of the neo-republican ideal of non-domination.

Section Two of this paper lays out Philip Pettit's account of the republican tradition, and in particular his understanding of liberty as non-domination, and its relation to his egalitarian theory of social justice. Section Three shows that Walzer has a very similar understanding of liberty as the opposite of domination. This section also points out how Walzer's concept of liberty relates to his political egalitarianism and how the latter differs from the simple egalitarianism he criticizes. Section Four explains Walzer's ideal of an autonomous distribution of all social goods and gives a preliminary analysis of his idea of "complex equality." Section Five demonstrates that Walzer conceives of complex equality as simple equality of liberty or freedom. In a "complex egalitarian society," every citizen is equally free from domination and tyranny. The conclusion contains a critique of both Walzer and Pettit. However, it argues that the two forms of republicanism are complementary because they each illuminate the other's blind spot and thus mutually fix each other's particular shortcoming.

2. *The Republican Understanding of the Relation of Liberty, Equality and Social Justice*

Philip Pettit understands his work as part of a growing number of contributions to political theory that is oriented towards a republican ideal or the republican tradition of thought.⁶ For Pettit, the recent republican movement began with Quentin Skinner's historical research on the Medieval foundations of modern political thought and his articles on Machiavelli, along with the works of others who write within a republican tradition identified as such by John Pocock (Pocock 1975, Skinner 1990: 293–309, 1998, 2002: 186–212. The latter essay also contains a critique of Isaiah Berlin's concept of negative liberty.). This tradition, which Pettit calls the "Italian-Atlantic," starts historically in Rome with Polybius, Cicero and Titus Livius. Leading Italian thinkers of the Middle Ages and the Renaissance—especially "the divine Machiavel'

⁵ In this book, Walzer explicitly sticks to the view he had already pronounced 1990 in his article "The Communitarian Critique of Liberalism." According to his summary of this view, communitarianism is not "a freestanding doctrine or substantive political program" but "a corrective to liberal theory and practice" (Walzer 2005a: x; cf. Walzer 2005b).

⁶ Cf. Pettit's list of contributions (Pettit 2012: 3 fn. 1).

of the *Discourses*”—base their work on these Roman scholars (Pettit 1997: 5, 2012: 6). In the Anglo-Saxon region, the republican tradition was continued by James Harrington, John Milton and Algernon Sidney, in France by Montesquieu.

Rousseau, on the contrary, constitutes, for Pettit, the beginning of the communitarian tradition that he regards as a continental form of republicanism in the broad sense. Pettit differentiates this form of republicanism from the Italian-Atlantic tradition (Pettit 2012: 11–12). On the one hand, Rousseau maintains one feature of the Italian-Atlantic tradition, the understanding of liberty as non-domination, while on the other, he breaks with this tradition by arguing against the two other features it holds to be characteristic for a republic: a mixed constitution and a contestatory citizenry:

The mixed constitution was meant to guarantee a rule of law—a constitutional order—under which each citizen would be equal with others and a separation and sharing of powers—a mixed order—that would deny control over the law to any one individual or body. The contestatory citizenry was the civic complement to this constitutional ideal: it was to be a citizenry committed to interrogating all the elements of government and imposing itself in the determination of law and policy. (Pettit 2012: 5)

In order to check the government and be vigilant, the citizens need individual and collective virtue. A contestatory citizenry and a functioning mixed constitution can protect all citizens from domination and, thus, protect their liberty (Pettit 2012: 5).

For Pettit, the new perspective that the republican tradition opens up on contemporary politics is mainly indebted to the innovative understanding of liberty as non-domination that goes back to Roman thought. However, this understanding was obscured by the liberal comprehension of liberty as non-interference, starting with the debates around the American Revolution. In line with this, republicanism was substituted by liberalism as the prevailing political philosophy (Pettit 1997: 12). Liberty as non-interference is identical with what, in *Two Concepts of Liberty*, Isaiah Berlin calls “negative liberty.” Berlin marks “negative liberty” off from “positive liberty” (Berlin 1958). Compared with this distinction, Pettit’s concept of liberty as non-domination, which goes back to Skinner, is as a third possibility.

For Berlin, “negative liberty” means the absence of interference—I am negatively free “to the degree to which no human being interferes with my activity” and to the degree that I can make uncoerced and unimpeded choices (Berlin 1958: 7). “Positive liberty” requires more than just not being disturbed or let alone by others, however; it means “self-mastery, with its suggestion of a man divided against himself” (Berlin 1958: 19)—I am positively free to the degree to which I achieve self-mastery, which is the rule of my better parts over my worse.

According to Pettit, the prevalence of Berlin’s distinction obscured the philosophical validity and historical reality of a third understanding of liberty: the republican concept of liberty as non-domination. This

understanding defines liberty also as absence, but as the absence of mastery or domination. The republican understanding is, so to speak, located in the middle between Berlin's alternative because it combines the focus on absence with the focus on mastery or domination. Pettit illustrates his comprehension of mastery and liberty with the relation of master and slave. The intellectual roots of this comprehension hark back to Roman thought (Wirszubski 1950).

The Romans conceived of the good of liberty (*libertas*) as the opposite of the unfree state of the slave (*dominatio, servitus*) who was subjected to his master (*dominus*). In the extreme case, the relation of master and slave means that the one who rules can interfere on an arbitrary basis with the choices of the other. Liberty as non-domination, however, equates to a state in which a person is more or less immune against the arbitrary interference of other persons. At the level of politics, this requires a political order that protects all citizens from the arbitrary interference of the powerful citizens and thus from domination. This is the republican ideal of social and political liberty (Pettit 1997: vii–viii, 22, 24, 27). Both of the republican ideas of a mixed constitution and of a contestatory citizenry serve to realize and to secure this ideal of freedom for all citizens.

In *On the People's Terms*, published in 2012, Pettit presents a republican theory of social justice. Like a republican state, such a theory has to be “substantively” and “expressively egalitarian,” and thus has to treat all citizens as equals (Pettit 2012: 81, 88, 297). In this context, Pettit refers to the egalitarian debate on the “Equality of What?” that has been ongoing for some decades. Should there be equality of income, of resources, of welfare, of rights, of capabilities to perform certain human functions or of some other aspect? (Pettit 2012: 77–81, 297).⁷ Like other egalitarians, Pettit singles out one form of equality that he considers to be the most relevant. His republican understanding of social justice aims at a society in which all citizens equally enjoy the status of being a free citizen. The concept of liberty that constitutes the republican ideal of “equal status freedom” defines liberty as the stable absence of domination: “A republican theory of justice would seek the expressive equalization of freedom as non-domination: the promotion of freedom as non-domination is the basis of an equal concern for each citizen” (Pettit 2012: 297, 123; cf. 298).

The realization of the republican understanding of social justice requires social institutions that safeguard equal liberty for all citizens. Such institutions comprise the provision of public resources for citizens, a high level of social security, and laws and norms. Pettit compares such institutions with antibodies in the bloodstream: just as antibod-

⁷ Cf. footnote 2. Sen explains: “While the question ‘why equality?’ is by no means dismissible, it is not the central issue that differentiates the standard theories, since they are all egalitarian in terms of some focal variable. The engaging question turns out to be ‘equality of what?’” (Sen 1992: 4. For an overview of the debate on *Equality of what?* cf. Cohen 1989).

ies relate to the immunity against certain diseases, so do institutions safeguard citizens from domination (Pettit 2012: 123–24, 128, 297–99). In order to determine “what level of support is sufficient” to efficiently render citizens immune to domination, Pettit introduces “the eyeball-test”: “They can look others in the eye without reason for the fear or deference that a power of interference might inspire; they can walk tall and assume the public status, objective and subjective, of being equal in this respect with the best” (Pettit 2012: 84). Citizens enjoy enough resources and safeguards if their extent is sufficient to pass the eyeball-test. For Pettit, a just and legitimate political order that meets the requirements of the republican ideal has to aim at a state that equally protects all citizens from domination and thus realizes their equal liberty.⁸

3. *Liberty or Domination?*

The concept of liberty plays a central role in John Rawls’s conception of justice as fairness. His first principle of justice, which Rawls gives priority over the second, calls for equal liberties: “Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all” (Rawls 1971: 302). By “equal basic liberties,” Rawls means classical individual or civil rights like liberty of conscience and freedom of thought and freedom of speech and assembly, as well as the democratic political rights of citizens to vote and to stand for public office (Rawls 1971: 61).

In *Spheres of Justice*, Walzer shows little interest in liberty in the sense of civil rights and liberties. This runs contrary to the approach of his previous book on *Just and Unjust Wars* (Walzer 2006),⁹ as well as to Rawls’s conception of justice. Walzer even considers the two most basic rights: not to be robbed of life or of liberty, “only of limited help in thinking about distributive justice” (Walzer 1983: xv). However, he declares that the kind of egalitarianism he is arguing for in his book “is consistent with liberty” (Walzer 1983: xiv). But what does Walzer mean by the liberty he associates with his political egalitarianism? A crucial quotation from the Preface of *Spheres* makes clear that Walzer has the same understanding of liberty as Skinner and Pettit:

The aim of political egalitarianism is a society free from domination. This is the lively hope named by the word *equality*: no more bowing and scraping,

⁸ Contrary to Rawls in *A Theory of Justice*, Pettit distinguishes between the terms “social justice” and “political legitimacy”. While social justice concerns the relations between people within the state, political legitimacy concerns the relations between “citizens as a whole and the state itself” (Pettit 2012: 130, 75).

⁹ In *Spheres*, Walzer states in retrospect: “Some years ago, when I wrote about war, I relied heavily on the idea of rights. For the theory of justice in war can indeed be generated from the two most basic and widely recognized rights of human beings—and in their simplest (negative) form: not to be robbed of life or of liberty” (Walzer 1983: xv).

fawning and toadying; no more fearful trembling; no more high-and-mightiness; no more masters, no more slaves. It is not a hope for the elimination of differences; we don't all have to be the same or have the same amounts of the same thing. Men and women are one another's equals (for all important moral and political purposes) when no one possesses or controls the means of domination. But the means of domination are differently constituted in different societies. Birth and blood, landed wealth, capital, education, divine grace, state power—all these have served at one time or another to enable some people to dominate others. Domination is always mediated by some set of social goods. (Walzer 1983: xiii, Walzer's italics)

The political egalitarianism Walzer is advocating doesn't aim at equalizing the unequal or giving everyone equal shares of some good. The kind of equality he strives for is a form of equality he characterizes as "free from every sort of domination" (Walzer 1983: xv). For Walzer, a society free from domination and subordination is not only a free society but also a society of equals. In such a society, citizens are equal in the sense that they are equally free from domination. Such a society would pass Pettit's eyeball-test because its citizens can indeed "walk tall" and "look others in the eye." For Walzer, a society in which all citizens are equally free from domination is also a just society. This is the central idea of the political philosophy he lays out in *Spheres*.

From Walzer's perspective, freedom and liberty are concepts opposed to "domination" and "dominance": a free society is a society free from domination and dominance. As such, these are key concepts in Walzer's theory of justice. For him, "dominance" is "the central issue in distributive justice" (Walzer 1983: 16). But what does Walzer mean exactly by the terms "dominance" and "domination," and how does he distinguish the two? Domination is, as quoted, "always mediated by some set of social goods," so insofar as these goods are dominant, they can serve as a means for domination. In our contemporary society, the main dominant social goods are money or capital, which are convertible into other social goods like an excellent education, recognition or political power and allow the rich to dominate the poor (Walzer 1983: 22, 315).

Walzer defines a dominant good in the following way: "I call a good dominant if the individuals who have it, because they have it, can command a wide range of other goods" (Walzer 1983: 10)¹⁰ a dominant good is easily converted into another good or into many others (Walzer 1983: 11). This illustrates why people try to accumulate dominant goods, like money. The central problem with dominance and dominant goods for Walzer is that "the dominance of goods makes for the domination of

¹⁰ Walzer goes on to say about the dominant good: "It is monopolized whenever a single man or woman, a monarch in the world of value—or a group of men and women, oligarchs—successfully hold it against all rivals" (Walzer 1983: 10–11). Walzer's term of a *monopoly* of social goods is somewhat misleading as in most societies it is rarely only *one* person who owns or controls some sort of social good.

people" (Walzer 1983: 19). Domination of people and its social prerequisite, the dominance of goods, is what Walzer considers to be unjust.

Of course, the problem of an unequal distribution of goods has often been addressed by the prevailing egalitarian political philosophers. In Walzer's terminology, these philosophers advocate "simple equality." Simple equality means numeric or arithmetic equality, which is realised if I have the same amount of a good as you do (Walzer 1983: 13–17).¹¹ Simple egalitarians hold that an unequal possession of the dominant good or the dominant goods is unjust. As a consequence, they advocate a more equal distribution and thus a redistribution of goods. The goal for these egalitarians is to establish simple equality, which is their understanding of a just distribution. The problem is only which kind of equality should be established.¹²

Walzer has a highly critical attitude of the concept of simple equality. He argues that a "regime of simple equality won't last for long" and "would require continual state intervention" (Walzer 1983: 14–15). Furthermore, simple equality is an extremely reductionist approach that is quite inappropriate for the plurality and complexity of our distributive systems and their issues. In *Spheres*, therefore, Walzer doesn't focus on how to get rid of simple inequalities but attends to possible ways and solutions to reduce dominance. The central claim of his theory of justice is that "the way should be opened for the autonomous distribution of all social goods: this amounts to saying that dominance is unjust" (Walzer 1983: 13). Walzer's goal is a society in which this "autonomous distribution" of all social goods is enforced, which, so he claims, prevents dominance.

4. *Autonomous Distributions and Complex Equality*

For Walzer, it is a fact that social goods tend to have different meanings in different societies. The claim of his interpretative method is that the proper distributive criteria of social goods are intrinsic to each particular social good. It is the meaning of each social good that determines the criterion of its just distribution.¹³ Walzer argues, for example, that the appropriate understanding of the meaning of medical care and welfare reveals to us that these goods should not be sold but allocated according to need (Walzer 1983: 64–91).

The consequence of Walzer's claim that the meaning of each social good determines its criterion of just distribution is that each social good and its distinct meaning constitutes—as he puts it metaphorically—a

¹¹ The concept of simple equality as numeric or arithmetic equality goes back to Plato and Aristotle (Aristotle, *Politics*, V 1 1301b 29–34; Plato, *Laws*, VI 757). Cf. the concept of simple equality Miller (1995b: 197–202).

¹² As mentioned previously, among egalitarians there has been a vivid debate going on about the crucial question: *Equality of What?*

¹³ Cf. for the difficulties of Walzer's claim that the meaning of each social good determines its criterion of just distribution Miller (1995a: 1–16, especially 5–10).

separate and relatively autonomous sphere of justice: "When meanings are distinct, distributions must be autonomous. Every social good or set of goods constitutes, as it were, a distributive sphere within which only certain criteria and arrangements are appropriate" (Walzer 1983: 10). In the case of medical care and welfare in general, these constitute a sphere in which the proper criterion for a just distribution is need. Public honour, on the other hand, constitutes a sphere in which the criterion is merit or desert, while office constitutes a sphere in which the suitable criterion is qualification (Walzer 1983: 135–139, 143–147, 259–262). And it is the distribution of these social goods according to their appropriate criteria that makes the distribution autonomous. Walzer is aware, though, that there is no absolute autonomy of the spheres but only a relative one, because what "happens in one distributive sphere affects what happens in the others" (Walzer 1983: 10). The autonomy is internal to the spheres, not in the system of their external, inter-sphere relations.

Contrary to an autonomous distribution, the allocation of medical care, public honour or offices to people who possess the dominant good of money means an "invasion" of these spheres and a violation of their inherent criteria: "Dominance describes a way of using social goods that isn't limited by their intrinsic meanings or that shapes those meanings to its own image" (Walzer 1983: 10–11). For Walzer, the latter equals a false interpretation and an overpowering of meaning. The result of Walzer's reflections on social goods and distribution is a general distributive principle which reads as follows:

The critique of dominance and domination points toward an open-ended distributive principle. *No social good x should be distributed to men and women who possess some other good y merely because they possess y and without regard to the meaning of x.* (Walzer 1983: 20, Walzer's italics)

Thus, in this formulation, the good y is a dominant good, and dominance would be to distribute x to people merely because they possess that dominant good y.¹⁴

In the preface of the German edition of his work on distributive justice, Walzer emphasizes that he holds the idea of complex equality to be the most interesting one of the book (Walzer 1992: 11). But again, what does "complex equality" mean exactly? Are there any good reasons to characterize Walzer's conception of just distributions as egalitarian in the sense of aiming at equality at all? If public honour is distributed according to desert or merit, for example, and welfare "in proportion to need," these distributions will lead to unequal results (Walzer 1983: 84). The same is true for higher education which, contrary to basic education, is to be distributed according to the criteria of interest and ca-

¹⁴ Indeed, according to Walzer there are three criteria for the distribution of goods that "meet the requirements of the open-ended principle": Free exchange, desert, and need (Walzer 1983: 21). Walzer claims "that every criterion that has any force at all meets the general rule within its own sphere, and not elsewhere" (Walzer 1983: 26). Cf. the problems of this claim Haus 2000, Den Hartogh 1999.

capacity (Walzer 1983: 203, 206). And even in the sphere of democratic politics Walzer sees no problem with an unequal distribution of political power.¹⁵

Walzer's idea of complex equality is his alternative to the notion of simple equality, for which egalitarians usually argue. What he means by complex equality is phrased in a central passage of his book:

The *regime of complex equality* is the *opposite of tyranny*. It establishes a set of relationships such that domination is impossible. In formal terms, complex equality means that no citizen's standing in one sphere or with regard to one social good can be undercut by his standing in some other sphere, with regard to some other good. Thus, citizen X may be chosen over citizen Y for political office, and then the two of them will be unequal in the sphere of politics. But they will *not be unequal generally* so long as X's office gives him no advantages over Y in any other sphere—superior medical care, access to better schools for his children, entrepreneurial opportunities, and so on. So long as office is not a dominant good, is not generally convertible, office holders will stand, or at least can stand, in a *relation of equality* to the men and women they govern. (Walzer 1983: 19–20, all italics by M. K.)

The opposition of the regime of complex equality to tyranny is crucial for an adequate understanding of Walzer's idea of complex equality—as indicated by the title of the final chapter of his book on distributive justice: *Tyrannies and Just Societies*. While Walzer holds a tyrannical society to be an unjust society, he regards a “complex egalitarian society” to be a just society.

Although Walzer uses the term “tyranny” in *Spheres* in a wide sense, he is well aware that the “immediate connotations of the word *tyrant* are political; its pejorative sense derives from centuries of oppression by chiefs and kings—and, more recently, by generals and dictators” (Walzer 1983: 282, Walzer's italics). For Walzer, in fact, every usage of political power that aims at getting access to goods from other spheres is tyrannical. Like dominance, tyranny in the most general sense comprises a disregard for the principles of justice internal to each distributive sphere and an aggressive entry so as to “invade” these spheres (Walzer 1983: 19, 315; cf. 10–11, 59). However, the original meaning of tyranny as an illegitimate and arbitrary rule over people is also central for Walzer's theory of justice. This connects him with republicanism, for which tyranny is the analogy of the rule of a master over a slave that has to be avoided by all means (cf. Saracino 2012). Analogously, Walzer's political theory in *Spheres* is a critique of tyrannical and therefore illegitimate and arbitrary rule that goes along with the domination of people and severe inequality and injustice. Indeed, one of his most important insights in *Spheres* is that an illegitimate rule over

¹⁵ Although he admits that everyone should have the right to “exercise minimal power” through the right to vote, he understands democratic politics as “a monopoly of politicians” and declares: “Democracy puts a premium on speech, persuasion, rhetorical skill. Ideally, the citizen who makes the most persuasive argument—that is, the argument that actually persuades the largest number of citizens—gets his way” (Walzer 1983: 309, 304; cf. 305).

goods is connected to an illegitimate rule over people. The latter can be derived from the former: “In political life—but more widely, too—the dominance of goods makes for the domination of people” (Walzer 1983: 19).

In a just society, inequalities within each sphere and even the monopolization of goods in one sphere are not necessarily inappropriate; different people succeed in different spheres, but there is no ruling class or all-round winner. Dominance and tyranny, on the other hand, are inimical to justice. In a complex egalitarian society, there is no dominant good that is convertible into other social goods and any convertibility is strictly limited. The people who hold political power guard the boundaries of the spheres and prevent powerful men and women from violating the appropriate standards of distribution and from “invading” the other spheres: “But we can only talk of a *regime* of complex equality when there are many boundaries to defend; and what the right number is cannot be specified” (Walzer 1983: 10, 28, Walzer's italics). Because in a complex egalitarian society no one possesses or controls the means of domination, it is a society free of tyranny, domination and subordination.

5. *Complex Equality as Simple Equality of Liberty*

In the literature, there is a controversy concerning the question of whether Walzer is a political egalitarian who tries to save the concept of equality by its reinterpretation or whether he is a non-egalitarian and a critic of the ideal of equality, as Angelika Krebs claims (Den Hartogh 1999, Haus 2012, 2014, Krebs 2012). Obviously, Walzer himself contributed to this disagreement because he criticizes the ideal of simple equality and yet advocates complex equality and gave his book the subtitle *A Defense of Pluralism and Equality*. This raises an essential question: what exactly does the term “equality” mean in Walzer's notion of a complex equality, or, what kind of equality is Walzer advocating with his call for its complex form?

Michael Haus distinguishes four egalitarian aspects in Walzer's theory of justice that clarify the intuition on which the conception of complex equality is based: equality as equal opportunities, equality through compensations (citizens “lose” in some spheres and “win” in others), equality as equal worth of all citizens (autonomous distributions allow no general ranking in society), and an equality of membership or citizenship in the political community (Haus 2000: 254–261, 2014: 40–44). Another interpretation is given by David Miller in his essay on “Complex Equality” in a volume of essays he edited together with Michael Walzer on *Spheres*. As Miller points out, we should look at complex equality “arising as a by-product of many separate distributions, each of which is in itself inequalitarian [...]. So here equality does not refer to the way some identifiable good is distributed, but describes

the overall character of a set of social relationships” (Miller 1995b: 198–199).¹⁶

According to Miller, therefore, “an overarching equality of status” is the “best interpretation” of complex social equality: “In a society which realizes complex equality, people enjoy a basic equality of status which overrides their unequal standing in particular spheres of justice such as money and power” (Miller 1995b: 199, 206). In equality of status, equal citizenship plays a cardinal role. All citizens are “enjoying an equal status *qua* citizens” (Miller 1995b: 206). According to Walzer, membership in some human community is the “primary good that we distribute to one another” (Walzer 1983: 31). Although the members of a community have the right to politically determine who they want to admit, they are morally constrained by the principle of mutual aid. Even more important, everyone who is admitted as a new immigrant, as a refugee or as a resident or worker “must be offered the opportunities of citizenship” (Walzer 1983: 62).

In his interpretation of complex equality as equality of status or equality of citizenship, Miller declares that the term “equality” should not be understood as a form of simple equality.¹⁷ However, the “kind of equality” manifested in the overall relationship of people in a complex egalitarian society has to be understood exactly as that. For Walzer, simple equality “is a simple distributive condition, so that if I have fourteen hats and you have fourteen hats, we are equal” (Walzer 1983: 18). By this definition, in a distribution of membership or status, all citizens are equal if they all have the same status of citizenship. In his reply to the critique of the egalitarian Richard Arneson, Walzer makes it clear that complex equality also appeals and goes back to simple equality: “complex equality is a version of equality; the adjective qualifies the noun, it doesn’t replace it” (Walzer 1995: 283, cf. Arneson 1995: 249–250, Walzer 2014: 9–14). In a complex egalitarian society people will be “in fact more equal, on some measure, than they are now” (Walzer 1995: 283).

For Walzer, the crucial measure, that on which people in a complex egalitarian society would be more equal, is freedom as non-domination. In his reply to Arneson, Walzer moves on to say that the egalitarianism of complex equality is “manifest in a radical decline of the dominance of some people over others” (Walzer 1995: 283). Such dominance—or,

¹⁶ From a perspective of complex equality, equality is not conceived as a “Zustand der Gleichverteilung bestimmter Güter, sondern als eine übergreifende Eigenschaft des gesellschaftlichen Zusammenlebens” (Haus 2003: 177).

¹⁷ Miller interprets “Walzer’s overarching notion of complex equality” as “the idea that in a society in which different people succeed in different spheres, their relationships overall can manifest a certain kind of equality. This is not simple equality, the sort that might obtain if people had equal amounts of property, or income. It is equality that comes about through many separate inequalities, cancelling or offsetting one another in such a way that no one can be picked out as an all-round winner” (Miller 1995a: 12).

more precisely, and as Walzer usually expresses it—such domination is “always mediated by some set of social goods”. Since the main dominant social good in contemporary societies is money, its exchange into other goods like education, welfare or political power has to be blocked (see Walzer’s list of all blocked exchanges Walzer 1983: 100–103). These are all goods that money should not be able to buy. How closely Walzer links the two concepts “equality” and “liberty” shows his conviction, that “it isn’t only equality but freedom, too, that we defend when we block a large number of (the larger number of) possible exchanges” (Walzer 1983: 317). As a formula: the less dominance of goods, the less domination of people over others and the more liberty, equality and justice. The most fundamental form of simple equality that is brought about in a society that achieves complex equality is not equality of status defined as equality of membership. On the contrary, the most crucial form of simple equality that Walzer anticipates as the result of autonomous distributions, and thus, of a complex egalitarian society, is equality of liberty or freedom, or, as Pettit puts it, “equal status freedom.”

In the preface to *Spheres*, Walzer understands his “entire book” as “an answer of a complicated sort” to the question: “In what respects are we one another’s equals?” (Walzer 1983: xii). His most fundamental answer to this question is: “Men and women are one another’s equals (for all important moral and political purposes) when no one possesses or controls the means of domination” (Walzer 1983: xiii). People are most importantly one another’s equals insofar as they are equally “free from every sort of domination” (Walzer 1983: xv, cf. 317). As, in a complex egalitarian society, everyone is equally free from domination, tyranny and subordination, complex equality should be interpreted primarily as simple equality of freedom or liberty. Thus a just and pluralistic society is also an equal and free society or a society of free and equal citizens. The concept of liberty and equality achieved by Walzer’s complex egalitarian society is primarily negative: it means the absence of tyranny or non-domination. This conception of liberty is the one that prevails in contemporary republican political thought.

Conclusion

Despite Walzer’s critique of simple equality, in the end his theory of justice aims primarily at simple equality of liberty or freedom. Though the “product of autonomous distributions” is, as Walzer declares, “complex equality,” it is also a simple equality of liberty or freedom (Walzer 1995: 283). However, this does not mean that Walzer contradicts himself or that his theory of justice is inconsistent. Although the state of equal freedom from domination is an essential feature of a complex egalitarian society, citizens in such a democratic society are equal and unequal in many different respects. They have unequal amounts of most social goods, inevitably, but they have “a single political status”

and there will be “equality of membership” (Walzer 1983: 62, 84). In order to ensure equal membership and avoid the exclusion of citizens, there is an equal basic education, an equal right to vote, “equal rights” in general, and suchlike (Walzer 1983: 202–203, 206, 305–306, 309). All these forms of equality are forms of simple equality, and, combined, they constitute complex equality or features of a complex egalitarian society. However, simple equality of liberty or freedom is the most important trait of such a democratic society.

Walzer criticizes political and theoretical approaches that aim at establishing simple equality by redistributing the dominant good for being reductionist and unstable. Certainly, the first of these two criticisms does not apply to a complex egalitarian society, in which all social goods are distributed according to their respective social meanings. Regarding the second, Walzer confesses that complex equality would not “necessarily be more stable than simple equality” (Walzer 1983: 17). His third criticism of a regime of simple equality is that it “would require continual state intervention” (Walzer 1983: 15). Although Walzer conceives of the state as the appropriate setting or framework of his whole theory of justice, however, he denies that this criticism applies to a complex egalitarian society (Walzer 1983: 28–30). He claims that, in such a society, “resistance to convertibility would be maintained, in large degree, by ordinary men and women within their own spheres of competence and control, without large-scale state action” (Walzer 1983: 17). However, he contradicts this claim by saying that it is political power that “is used to defend the boundaries of all the distributive spheres, including its own, and to enforce the common understandings of what goods are and what they are for” (Walzer 1983: 15 fn., 281).¹⁸ Nevertheless, Walzer admits that political power can be used “to invade the different spheres and to override those understandings” (Walzer 1983: 15 fn.).

Although Walzer devotes a long chapter of *Spheres* to political power, he does not explain how office holders can be prevented from invading the different spheres. There is another serious problem that Walzer does not address. Even if all goods were distributed according to their social meanings, the people who hold political power could still possess the power to interfere in the affairs of the citizens they govern on an arbitrary basis. These two problems represent a blind spot of Walzer's theory of justice: essentially, it offers elements of a theory of democracy but no theory of government. Pettit's neo-republicanism, however, is able to fix this shortcoming. As the subtitle of Pettit's book *Republicanism* indicates, he offers not only a theory of freedom but of government (Pettit 1997). Pettit's version of republicanism devotes considerable attention to the question of how to check the government or the people

¹⁸ For Walzer, political power is “probably the most important, and certainly the most dangerous, good in human history” (Walzer 1983: 15). Cf. the chapter on political power (Walzer 1983: 281–311).

who hold political power. For him, a contestatory citizenry and a functioning mixed constitution can protect all citizens from domination.

However, Pettit's neo-republicanism has its own blind spot. Like Skinner, Pettit conceives of domination primarily as a direct relation among people. Domination can occur in the workplace and in the family as domination between private individuals. However, domination is also mediated through political institutions. Government power is supposed to be a remedy for the domination between private individuals, but can lead to a new form of domination that occurs between the holders of political power and the people they govern. Although Pettit is right to conceive of domination as a direct relation among people, he neglects to consider the fact that domination is usually also mediated through social goods. Contrary to Walzer, Pettit has no theory of social goods. It is a main strength of Walzer's theory that it focuses on the different social goods through which domination is mediated and that he is able to demonstrate that dominant goods serve as a means for the domination of people. Therefore, Walzer's theory is able to fix an important shortcoming of Pettit's theory. As a result, Pettit and Walzer pick out different means and ways how to reach the republican goal at which they both aim at, the freedom of citizens from domination; each succeeds in a place where the other fails, and so, their two republican theories should be viewed as allies that complement each other.

References

- Arneson, R. J. 1995. "Against 'Complex' Equality." In Miller and Walzer 1995: 226–252.
- Avineri, S. and de-Shalit, A. (eds.) 1992. *Communitarianism and Individualism*. Oxford-New York: Oxford University Press.
- Benbaji, Y. and Sussmann, N. (eds.). 2014. *Reading Walzer*. Abingdon: Routledge.
- Berlin, I. 1958. *Two Concepts of Liberty*. Oxford: Oxford University Press.
- Cohen, G. A. 1989. "On the Currency of Egalitarian Justice." *Ethics* 99: 906–944.
- Den Hartogh, G. 1999. "The Architectonic of Michael Walzer's Theory of Justice." *Political Theory* 27 (4): 491–522.
- Dworkin, R. 1981a. "What is Equality? Part 1: Equality of Welfare." *Philosophy and Public Affairs* 10: 181–246.
- Dworkin, R. 1981b. "What is Equality? Part 2: Equality of Resources." *Philosophy and Public Affairs* 10: 283–345.
- Dworkin, R. 2001. *Sovereign Virtue. The Theory and Practice of Equality* (Third Printing), Cambridge: Harvard University Press.
- Haus, M. 2000. *Die politische Philosophie Michael Walzers. Kritik, Gemeinsinn, Gerechtigkeit*. Wiesbaden.
- Haus, M. 2003. *Kommunitarismus*. Wiesbaden.
- Haus, M. 2012. "Blockierter Tausch—befreite Gerechtigkeit? Walzer und die Tyrannei des Geldes." In K-H. Nusser (ed.). *Freiheit, soziale Güter*

- und Gerechtigkeit. Michael Walzers Staats- und Gesellschaftsverständnis.* Baden Baden: 59–88.
- Haus, M. 2014. "Komplexe Gleichheit—Wie egalitär ist Walzer's Gerechtigkeitstheorie." In M. Knoll and M. Spieker (eds.) *Sphären der Gerechtigkeit. Ein kooperativer Kommentar.* Stuttgart: 27–50.
- Krebs, A. 2012. "Gleichheit und Gerechtigkeit. Wie egalitaristisch ist Michael Walzers Ideal der 'komplexen Gleichheit'?" In K-H. Nusser (ed.). *Freiheit, soziale Güter und Gerechtigkeit. Michael Walzers Staats- und Gesellschaftsverständnis.* Baden Baden: 89–104.
- Kymlicka, W. 2002. *Contemporary Political Philosophy. An Introduction.* Second Edition. Oxford: Oxford University Press.
- Miller, D. and Walzer, M. (eds.). 1995. *Pluralism, Justice and Equality.* Oxford: Oxford University Press.
- Miller, D. 1995a. "Introduction." In: Miller and Walzer 1995: 1–16.
- Miller, D. 1995b. "Complex Equality." In: Miller and Walzer 1995: 197–225.
- Nusser, K-H. (ed.). 1995. *Freiheit, soziale Güter und Gerechtigkeit. Michael Walzers Staats- und Gesellschaftsverständnis.* Baden Baden.
- Pettit, P. 1997. *Republicanism. A Theory of Freedom and Government.* Oxford: Oxford University Press.
- Pettit, P. 2012. *On the People's Terms. A Republican Theory and Model of Democracy.* Cambridge: Cambridge University Press.
- Pocock, J. 1975. *The Machiavellian Moment: Florentine Political Theory and the Atlantic Republican Tradition.* Princeton: Princeton University Press.
- Rawls, J. 1971. *A Theory of Justice.* Cambridge: Harvard University Press.
- Rawls, J. 1999. *A Theory of Justice.* Revised Edition. Cambridge: Harvard University Press.
- Reiner, T. 2011. "The sources of communitarianism on the American left: Pluralism, republicanism, and participatory democracy." *History of European Ideas* 37: 293–303.
- Saracino, S. 2012. *Tyrannis und Tyrannenmord bei Machiavelli. Zur Genese einer antitraditionellen Auffassung politischer Gewalt, politischer Ordnung und Herrschaftsmoral,* Humanistische Bibliothek, Vol. 62. München.
- Sen, A. 1992. *Inequality Reexamined.* Oxford: Oxford University Press.
- Skinner, Q. 1978. *The Foundations of Modern Political Thought.* Cambridge: Cambridge University Press.
- Skinner, Q. 1990. "The Republican Ideal of Political Liberty." In G. Bock, Q. Skinner and M. Viroli (eds.). *Machiavelli and Republicanism.* Cambridge: Cambridge University Press: 293–309.
- Skinner, Q. 1998. *Liberty before Liberalism.* Cambridge: Cambridge University Press.
- Skinner, Q. 2002. "The Idea of Negative Liberty: Machiavelli and Modern Perspectives." In Q. Skinner, *Visions of Politics, Vol. 2: Renaissance Virtues,* Cambridge.
- Swift, A. 1995. "The Sociology of Complex Equality." Miller and Walzer 1995: 253–280.
- Walzer, M. 1980. *Radical Principles: Reflections of an Unreconstructed Democrat.* New York: Basic Books.

- Walzer, M. 1983. *Spheres of Justice. A Defense of Pluralism and Equality*. New York: Basic Books.
- Walzer, M. 1992. *Sphären der Gerechtigkeit. Ein Plädoyer für Pluralität und Gleichheit*. Frankfurt am Main-New York: Campus Verlag.
- Walzer, M. 1995. "Response." In Miller and Walzer 1995: 281–297.
- Walzer, M. 2005a. *Politics and Passion. Toward a More Egalitarian Liberalism*. New Haven-New York: Yale University Press.
- Walzer, M. 2005b. "The Communitarian Critique of Liberalism." In Walzer 2005a : 141–163 (first in: *Political Theory* 18 (1): 6–23).
- Walzer, M. 2006. *Just and Unjust Wars. A Moral Argument with Historical Illustrations. Fourth Edition with a New Introduction by the Author* (Orig. 1977). New York: Basic Books.
- Walzer, M. 2007a. *Thinking Politically: Essays in Political Theory*. Selected, edited, and with an introduction by David Miller. New Haven: Yale University Press.
- Walzer, M. 2007b: "The Civil Society Argument. A Path to Social Reconstruction." In Walzer 2007a.
- Wirszubski, C. 1950. *Libertas as a Political Idea at Rome during the Late Republic and Early Principate*. Cambridge: Cambridge University Press.

*The Epistemic Value of Partisanship**

IVAN CEROVAC
University of Rijeka, Rijeka, Croatia

This paper discusses the epistemic value of political parties and other partisan associations from the standpoint of epistemic democracy. It examines whether political parties contribute to the quality of democratic deliberation, thus increasing the epistemic value of democratic decision-making procedures, or represent a threat that polarizes the society and impedes and distorts the public deliberation. The paper introduces several arguments that support the epistemic value of partisanship. Partisan associations empower otherwise marginalized social groups or groups that have disproportionately small political influence by facilitating political education or by connecting citizens and experts who share the same values. Partisan associations also help us resist the epistemically damaging effects of hermeneutical (epistemic) injustice by enabling marginalized citizens to construct alternative discourses. However, though partisanship might facilitate the transmission of knowledge, this deliberative tool will only be used in a group of like-minded citizens (i.e. within a political party), thus increasing the polarization between the parties and citizens alike, and decreasing the epistemic value of such collective decision-making procedures. The paper analyses some epistemic strategies (like red-teaming or building a critical thinking culture) that can help us avoid or (at least) reduce the epistemically damaging effects of polarization. However, internal action (from within a deliberative group) might not be enough. Making the deliberation on political issues public and spreading it through different forms of citizens' organizations will ensure that political deliberation is not closed within a single homogenous deliberating group (i.e. the party). These practices should significantly reduce the damaging effects of group polarization.

Keywords: Epistemic democracy, partisanship, group polarisation, affiliation, epistemic injustice.

* This paper is produced within the project "Well-being, affiliation and social justice" (UIP-2017-05-3462) co-financed by the Croatian Science Foundation.

Political parties have been studied for centuries, yet not much attention was brought to their epistemic value. In fact, political parties have often been thought of as subversive elements that endanger the epistemic qualities of democratic decision-making procedures: they organize citizens in order to promote the private, but not the public good (Rousseau 1997), they have a damaging effect on their members' epistemic capacities (Atchison 2012), they polarize the democratic society (Layman, Carsey and Horowitz 2006) and even lead to destimulation of voter turnout (Brady, Ferejohn and Harbridge 2008). It is therefore very interesting to study the value of political parties from the standpoint of epistemic democracy, a theory of democracy that ascribes to democracy its legitimacy-generating potential (at least in part) in virtue of its ability to produce (substantively) good or correct outcomes. The ability of the decision-making procedure to produce legitimate decisions is evaluated, in part, in regard to its ability to produce outcomes that are right, true or correct according to some procedure-independent standard. We evaluate political institutions and organizations (at least in part) by their ability to contribute to the procedure's substantive epistemic value, i.e. its ability to produce correct decisions.¹ Having this definition in mind, a few questions arise. Do political parties contribute to the quality of democratic deliberation, thus increasing the epistemic value of democratic decision-making procedures,² together with the legitimacy-generating potential of such procedures? Or is the opposite true that political parties polarize the society, impede and distort the public deliberation and damage the legitimacy-generating potential of democratic decision-making procedures?

This paper tries to answer the abovementioned questions by combining the traditional standpoint of normative political theory with some contemporary methods and models from social epistemology and group psychology. It is heavily influenced by Jonathan White's and Lea Ypi's *The Meaning of Partisanship*, though it expands well beyond the scope of the book, combining their approach to political partisanship with the epistemic account of democratic legitimacy. It aims to support and expand White and Ypi's idea that political parties increase the epistemic value of democratic procedures (though, it seems, using a different approach to political legitimacy), but also to specify conditions in which partisanship can have this epistemic value. It is important to emphasize that, even if this paper fails to convince that partisanship

¹ The epistemic approach to democracy was first formulated by Joshua Cohen (1986). The formulation used in this paper follows a (somewhat simplified version of) David Estlund's (2008) account of epistemic democracy.

² This paper does not differentiate strictly between collective decision-making procedures and procedures of collective authorisation of decisions. The account of epistemic democracy presented in this paper is thus compatible with various models of representative democracy, as well as with the recognition of experts in politics. For a detailed account on the distinction between these two types of procedures see Pavićević and Simendić (2016).

has epistemic value (or if it turns out to reduce the epistemic quality of democratic decision-making procedures), it might still have other (moral or political) values that support it.

The *first* part of the paper discusses new arguments for the epistemic contribution of partisanship presented by White and Ypi. Organizing in political parties can empower otherwise marginalized social groups or groups that have disproportionately small political influence by facilitating political education or by connecting citizens and experts who share the same values. Furthermore, partisanship can help us resist the epistemically damaging effects of hermeneutical (epistemic) injustice by enabling marginalized citizens to construct alternative discourses. The *second* part introduces some new arguments in support of the epistemic value of partisanship. These arguments highlight that partisanship can be a valuable tool for knowledge transmission and as such could represent a valuable asset for deliberative democracy. The danger of group polarization and crippled epistemology is reintroduced in the *third* part of the paper. Though partisanship might facilitate the transmission of knowledge, this deliberative tool will only be used in a group of like-minded citizens (i.e. within a political party), thus increasing the polarization between the parties and citizens alike. This paper uses results from contemporary empirical studies that suggest there is a widening gap between political parties in United States due to group polarization and crippled epistemology while also analyzing the damage these processes cause to the epistemic value of collective decision-making procedures. Some possible solutions are discussed in the *fourth* and the *fifth* part of the paper. Namely, a deliberative group can use some epistemic strategies (like red-teaming or building a critical thinking culture) to avoid or (at least) reduce the epistemically damaging effects of polarization. However, internal action (from within a deliberative group) might not be enough. Making the deliberation on political issues public and spreading it through different forms of citizens' organizations will ensure that political deliberation is not closed within a single homogenous deliberating group (i.e. the party). These practices should not only prevent crippled epistemology, but also significantly reduce the damaging effects of group polarization.

1. *White and Ypi on epistemic value of partisanship*

Jonathan White and Lea Ypi (2016) do not take the standard account of epistemic democracy as a broader normative framework for their work. Their approach follows and builds upon Bohman's (1998) position, which shifts focus from the quality of the results of the deliberative procedure to the quality of the procedure itself. Their theory is settled in the framework where the focus is "shifted from the outcome of justificatory practices to the normative significance of the process" (White and Ypi 2011: 392, 2016: 73). They measure the epistemic value of partisanship by examining its influence on the (intrinsic) moral and epistemic

quality of the collective decision-making procedure.³ Furthermore, they indicate and describe two mechanisms by which partisanship increases the epistemic value of the decision-making procedures.

First, partisanship can reduce the impact of power asymmetries on agents' capacities to participate as equals in reason-giving and decision-making processes. Though formal equality of political influence represents one of the most fundamental values of democratic societies, equality is rarely achieved (and sometimes deemed unnecessary or even adverse) in the informal political sphere.⁴ Social and economic inequalities often spill over to the informal political sphere thus making some groups unable to participate as equals in democratic procedures. Members of powerless groups will be disadvantaged not only in regard to the access to political education and specialized (expert) knowledge but also in access to the resources needed to convey their political message to the wider public. These inequalities in political influence impair both the moral and the epistemic value of democratic procedures thus endangering their legitimacy-generating potential. Partisanship helps us preserve that by cutting the link between social and economic power, and giving access to political education and expert knowledge. Ypi and White hold that partisanship plays this role by offering certain irreplaceable epistemic resources. Partisan forums (including party conventions, branch meetings, assemblies, protests, blogs and websites) can then be seen as learning platforms for citizens. They empower disadvantaged citizens and give them epistemic resources needed for political participation. Furthermore, partisan forums have an important motivational role—they show disadvantaged citizens that they are not alone in their political struggle. Therefore, partisanship “plays an important role in ensuring the sustainability of shared political projects when epistemic challenges are at stake” (White and Ypi 2016: 90).

The epistemic value of partisanship is first described through its educational role. Citizens who lack the access to education in general, or to political education in particular, gain new information and skills to make their political views more coherent, appealing and clearer to the wider audience. Etienne Lantier, the protagonist of Zola's *Germinal*, is taken to be a great example of such political education: he starts as a poorly educated, rebellious and unemployed young man who through his participation in epistemically-enriching partisan associations turns into an intellectually sophisticated activist (Zola 1983, as

³ Fabienne Peter (2009) defended a similar position, arguing that legitimacy-generating potential of collective decision-making procedures rests in their non-instrumental epistemic qualities. Some arguments against this view can be found in Marti (2006), Estlund (2008) and Cerovac (2016).

⁴ I use Estlund's (2008) differentiation between formal political sphere, which includes voting procedures and public elections, but also political institutions such as courts and legislators, and informal political sphere, which includes political speeches, candidate and citizen debates, opinion journalism, political advertising, political art and demonstrations.

paraphrased in White and Ypi 2016). Political parties and related associations (political foundations, think tanks, informal groups or even trade unions) can be great platforms for systemizing and spreading political knowledge, and for improving some relevant skills (verbal and non-verbal communication, networking, management and leadership skills). The second epistemic benefit of partisanship, often closely connected to the first one, is its ability to connect citizens and experts who otherwise would not come into contact. Since partisan associations (unlike factions) gather around some central values (and not around private interests of group members), they often include members from various backgrounds and fields of expertise. Their joint political effort and focus on the same political aims, combined with partisan forums as means for inter-party deliberation and learning, ensures that complex views and topics requiring technical knowledge can become available to all citizens.

Second, partisanship can help marginalized and disadvantaged groups to develop hermeneutic resilience. Namely, powerless groups are often unable to participate in the creation of collective hermeneutical resources which leaves them unable to make sense of their own social experiences (or at least unable to formulate them in a way other citizens could understand). This is characterized as hermeneutical epistemic injustice (Fricker 2007, 2013), and is often considered one of the effects that perpetuate social inequalities by keeping the members of disadvantaged groups from participating as equals in the process of collective decision-making.⁵ For example, terms like ‘sexual harassment’ and ‘post-natal depression’ have been created some fifty years ago, though such practices and occurrences have been around for thousands of years. They have affected women for centuries yet until recently women were unable to formulate and explain what is exactly going on (since there were no available hermeneutical resources to differentiate between sexual harassment and harmless flirting). These instances of hermeneutic injustice were removed by organized group action, when members of a disadvantaged group came together, deliberated their social experiences and formulated plans for social engagement. Ypi and White believe that partisanship associations can help us remedy hermeneutic (epistemic) injustice by creating new platforms and channels of communication. Namely, partisanship enables citizens to construct alternative discourses, to exchange the information they were unable to exchange before and to raise consciousness of the problematic aspects of common-sense thinking.

Partisanship is valuable since it offers certain irreplaceable epistemic resources: new channels and platforms of communication that enable citizens to express their social experiences and new resources

⁵ Martha Nussbaum (2000) and Amartya Sen (2003) defend similar ideas by emphasizing how adaptive preferences can keep members of disadvantaged groups from formulating and even desiring policies that might improve their social and economic status.

needed to process the information in a particular way, including interpretative resources needed for upholding partisan commitment. White and Ypi (2016: 94) describe a hypothetical case of Rosa, a socialist living in West Europe in 1989. She suddenly has to (in light of the new evidence) decide whether to completely abandon the idea of socialism or to revise her ideas and strengthen them with what can be learned from the collapse of socialist regimes in East Europe. It is very difficult for her to keep both partisan commitment and her faith in socialism when she is unable to deliberate with her fellow associates. Political parties and other partisan associations can help by structuring such deliberation and offering hermeneutical resources (but also expert knowledge) needed to uphold the partisan commitment.

White and Ypi seem to be aware that collective deliberation with like-minded citizens (epistemic trustees) can lead to epistemically damaging results, including group polarization and the creation of echo-chambers (Sunstein 2009). However, since their position does not evaluate the quality of collective decision-making procedure in its ability to achieve some epistemically valuable (true, correct or justified) results but by intrinsic (purely procedural) qualities of a decision-making procedure, they do not seem to be affected by these objections. In fact, they are ready to accept that deliberation within partisan organizations often does not lead to epistemically best (or even decent) outcomes:

If there is a truth of the matter to be found, exposure to disagreement and to the discursive challenge of one's own position may be more likely to contribute to an enlightened assessment. Precisely because arguing with one's political friends is more likely to consolidate one's previous opinions and beliefs and develop resistance to the tendency to revisit such commitment too lightly, taking part in associative practices strengthens the ability to stick with previously held beliefs and values, despite evidence that would suggest the need to revisit them. (White and Ypi 2016: 95)

Epistemic value of partisan associations is therefore not outcome-oriented—White and Ypi do not think that the role of partisanship (and its epistemic value) is to identify political projects worth pursuing but instead to provide means needed for sustaining them under epistemic pressure.

My view, which builds on the standard (outcome-oriented) account of epistemic democracy, is directly targeted by Sunstein's arguments on group polarization and echo-chambers. This paper does not deny the importance of these arguments—it instead builds upon them and tries to determine whether there are some (instrumental, and not intrinsic) epistemic qualities of partisan associations, as well as whether partisan associations can be (at least partly) protected from epistemic deviations.

2. *New arguments supporting the epistemic value of partisanship*

Although White and Ypi claim that partisanship has (only) intrinsic epistemic value, their arguments can also be used to extend the claim to instrumental epistemic value. Epistemic resources (e.g. partisan forums) that reduce the impact of power asymmetries on the political influence of the citizens by providing them political education can help us reach epistemically better outcomes. Furthermore, partisan associations can help marginalized and disadvantaged groups develop hermeneutic resilience, thus enabling them to make sense of their social experiences and to participate in decision-making processes that aim at producing better laws, policies and decisions. White and Ypi have thus (perhaps unintentionally) presented initial arguments supporting the instrumental epistemic value of partisanship.

This part of the paper analyses some positive effects of partisanship on knowledge transmission, building upon White and Ypi's account of partisanship associations that provide some epistemic resources that other kinds of associations would be unable to deliver. Namely, although deliberating with like-minded citizens often leads to some forms of 'crippled epistemology' (Sunstein 2009), it can also have some positive epistemic effects, primarily in the area of knowledge transmission from experts to citizens.

Democratic procedures have often been criticized for their inability to adequately incorporate specialized and expert knowledge in the decision-making processes. Experts⁶ often withdraw from public deliberation thus leaving those who remain engaged in the deliberative process to face extreme difficulties by trying to get the necessary trust from the citizens, and in making their arguments available for those who lack the specialized knowledge (Solomon 2006, Prijić-Samaržija 2017). Can intra-party deliberation represent a better institutional arrangement than public deliberation when it comes to employment of expert knowledge in collective decision-making processes? This paper proceeds by claiming that there are a few important aspects of partisanship that make deliberation conducted within such associations epistemically better than the one conducted within non-partisan associations or within general public. Unlike some non-partisan associations (or the public), political parties are characterized by epistemically relevant features such as solidarity, overlapping understanding, competition and sanctions⁷ (Christiano 2012). These features can improve

⁶ Discussion on experts and expertism in contemporary epistemology is very broad and diverse. This paper endorses a widely accepted definition by Adam Elga (2007), who claims that experts are those who are considerably less likely to have incorrect belief regarding certain issue. Experts are therefore seen simply as 'not-epistemic-peers'.

⁷ This does not imply that political parties are the only type of association characterized by these features. Some religious associations, for example, might

the knowledge transmission from experts to citizens, thus improving both the intrinsic and the instrumental epistemic quality of a decision-making process.

First, consider the positive effects solidarity has on the process of knowledge transmission. Accepting the testimony of others, even if they are experts, calls for assessment of their trustworthiness (Prijić-Samaržija 2011). This can be very difficult to achieve since both technical expertise and values held by experts might be inaccessible to regular citizens. Partisans, however, know that their partisan experts hold the same goals, values and ideals as themselves. Furthermore, partisans can see that other members of their party, those more competent in the relevant fields, recognize and praise the technical expertise of the experts in question, and are therefore more prone to trust them even when their opportunities and capacities for monitoring experts are limited. This is of crucial importance when experts have technical knowledge that is difficult or almost impossible to explain to non-experts. Expert knowledge cannot have the appropriate weight in the (democratic) decision-making process unless the experts are backed-up by non-experts who do not have the complete grasp of the knowledge in question but nonetheless defer to the (epistemic) authority of experts. This deference is facilitated by partisan solidarity.

Second, since parties are plural associations and have members with expertise in various areas, they can achieve a form of overlapping understanding among members. This becomes useful when two or more persons share some expertise but do not share other expertise. For example, A knows about disciplines *a*, *b* and *c*, while B knows about *b*, *c* and *d*—they overlap at *b* and *c*, and can translate some of their knowledge about *a* or *d* through ideas of *b* and *c* (Christiano 2012). Of course, some knowledge will be lost or simply cannot be properly transferred through overlapping understanding. This depends on the sophistication of theories in question, but also on the proximity of disciplines. It will be much easier for a political scientist to transfer his knowledge using overlapping understanding to an economist than to a physicist. However, overlapping understanding enables the transfer of specialized knowledge from two or more agents to the general public. A complex economic theory can be transferred to a political scientist, who can then combine it with his knowledge of the legal and political background and transfer it onwards to politicians or perhaps to relatively sophisticated journalists (Christiano 2012: 39–40). Politicians could then create laws and public policies based on (or at least taking into consideration) that specialized expert knowledge, and journalists could explain what they understand to ordinary citizens. Parties are composed of people proficient in different disciplines and gather experts in various fields, thus enabling the transfer of knowledge through overlapping understanding.

exhibit some or even all of the epistemic features described above.

Third, unlike many civil society associations, political parties directly compete for power and influence. The competition between parties fosters the competition between different political programs in the public arena, with each party trying to build a strong argument for its program and to find mistakes in the justification of other parties' programs. Parties have to rely on the knowledge of experts both to produce the program that advances their core values properly and to recognize when another party's program rests on mistakes that can be scientifically exposed. The competition between political parties has an epistemic (though not only epistemic) dimension, and it can thus increase the epistemic quality of political decision-making process.

Fourth, political parties have mechanisms for monitoring and evaluation of the work done by experts and for sanctioning experts when they depart from party's core values. Just like academic community can sanction scientists for methodological issues, political parties can sanction partisan-experts who betray the values of the party. This simultaneously builds intraparty solidarity by making it easier for party members to trust party experts and makes the system more democratic since it prevents (or at least discourages) experts from abandoning values supported by party members and the party in general.

To sum up, the initial argument by White and Ypi, supported with Christiano's position on positive effects of partisanship on knowledge transmission, shows that *some aspects* of partisanship (done properly) promote both the epistemic and the moral value of democratic decision-making procedures. Partisanship makes procedures fairer by removing some obstacles in the informal political public sphere that endanger fairness of the decision-making process. It also makes procedures more epistemically reliable by removing damaging effects of hermeneutical epistemic injustice and by enabling transference of knowledge between experts and non-experts. However, the question remains whether these epistemically positive aspects of partisanship exceed the damage it can cause to the collective decision-making process.

3. *Group polarization and crippled epistemology*

The problem of group polarization represents a permanent challenge for all kinds of epistemic democracy, including both purely procedural accounts that aim for some intrinsic epistemic values inherent in the decision-making procedure and rational procedural accounts that aim for substantive quality of political outcomes. Cass Sunstein (2005) describes group polarization as a statistical regularity when members of deliberative group embrace more extreme views than the ones they've had before the deliberation. Various people holding similar views thus, after the deliberation with like-minded citizens, end up holding more extreme versions of these views. For example, a group of moderate feminists (or nationalists or liberals) will, after repeated inter-group deliberation, become a group of extreme feminists (or nationalist or

liberals). The shift can be explained by every member of a group changing his or her views through deliberation with like-minded epistemic peers. Group polarization therefore presents a serious epistemic problem since citizens' views are changed only in one direction (towards the more extreme version) regardless of any reasons for or against such a view.

The effect is much stronger in 'deliberative enclaves' (Sunstein 2002, 2005, 2009), groups of like-minded persons deliberating on a regular basis for a longer period of time. However, political parties can also be described as 'groups of like-minded persons deliberating on a regular basis for a longer period of time'. After all, those are some of the key components of political parties: they (i) are composed of like-minded or similar-minded citizens, they (ii) encourage and foster deliberation between party members, and finally, they (iii) are long term political projects persisting for longer periods of time (unlike protest groups or civil initiatives). Political parties might thus be seen as a specific form of deliberative enclaves. This raises doubts regarding the epistemic value of political parties and the negative impact they might have on collective decision-making processes.

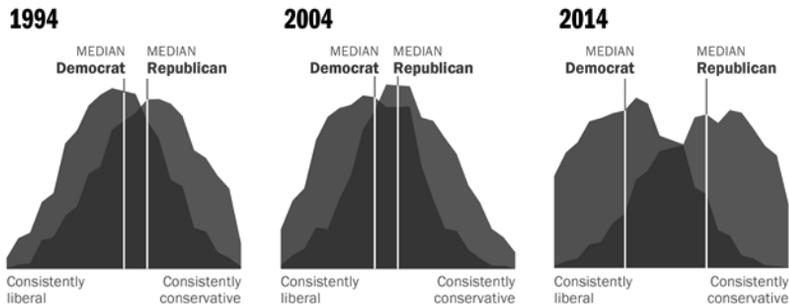
Group polarization leads to more extreme political views and extreme views require extreme cognitive conditions to be protected and nourished. This often leads to 'crippled epistemology' (Hardin 2002), another dangerous effect that takes place when, in order to preserve extreme views from challenges, groups adopt rules that discourage or even sanction deliberation with persons of different (political) views. Namely, since extreme views produced by group polarization are not grounded in good epistemic reasons (but are often created despite them), it is very difficult to continue defending them in open deliberation with citizens who do not share such views and introduce new reasons and arguments in the public debate. A simple method for dealing with such epistemic challenges is *argumentum ad hominem*, a fallacious argumentative strategy that "attacks the characteristic or authority of the author without addressing the substance of the argument" (Graham 2008). Members of extreme groups thus often employ this argumentative strategy to avoid criticism or objections to their views while deliberative groups often encourage and promote such strategies. Political adversaries and citizens who do not share the same political views as members of deliberative enclaves are characterized as stupid or malevolent (or both). This removes or weakens the demand to answer to objections and critiques from such citizens, thus keeping extreme views safe from possible challenges. This kind of discourse is becoming dominant in contemporary politics, with prominent examples like Ann Coulter (2004, as cited in Talisse 2009), a conservative commentator who advises her readers to talk to liberals only "if you must", and Michael Savage who claims that his liberal opponents suffer from "mental disorder" (Savage 2005, as cited in Talisse 2009) or Al Franken (2004)

characterizing republicans as “lying liars” and “stupid white men”.

Ideological division between political parties has been growing rapidly for the past fifteen years. The rise of new populist, extreme far-right and far-left parties in Europe supports this assumption, though it makes empirical research more difficult since it is no longer possible to track political attitudes of party members through longer periods of time (since some of these parties are only a few years old). However, empirical data from the United States, where political attitudes of supporters of the two dominant parties have been tracked for decades, indicates the growing gap between liberals and conservatives (Pew Research Centre 2014).

Democrats and Republicans More Ideologically Divided than in the Past

Distribution of Democrats and Republicans on a 10-item scale of political values



Source: 2014 Political Polarization in the American Public

Notes: Ideological consistency based on a scale of 10 political values questions (see Appendix A). The blue area in this chart represents the ideological distribution of Democrats; the red area of Republicans. The overlap of these two distributions is shaded purple. Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

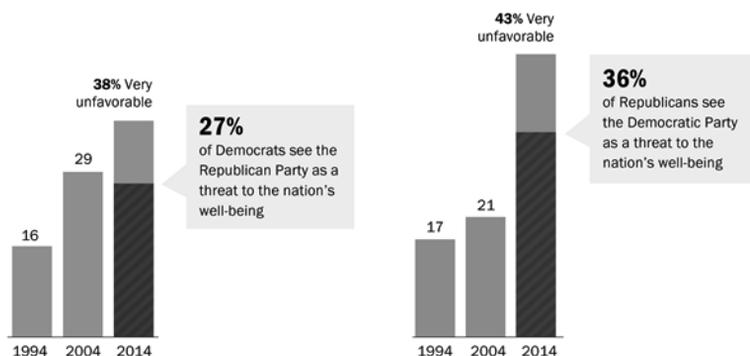
PEW RESEARCH CENTER

The growing gap between political attitudes of liberals and republicans need not be a problem in itself, however, it brings along many epistemically damaging features that simultaneously enhance group polarization and reduce social trust between citizens supporting different political parties. Group polarization is increased when political values and attitudes descend to the private sphere, with growing number of liberals rejecting to marry someone who is a gun owner and growing number of conservatives refusing to marry someone who is of different race or is born and raised outside the US (Pew Research Centre 2014). Social trust is reduced with increasing numbers of citizens seeing their political opponents not merely as someone they disagree with, but also as a potential threat to the nation’s well-being.

Beyond Dislike: Viewing the Other Party as a 'Threat to the Nation's Well-Being'

Democratic attitudes about the Republican Party

Republican attitudes about the Democratic Party



Source: 2014 Political Polarization in the American Public

Notes: Questions about whether the Republican and Democratic Parties are a threat to the nation's well being asked only in 2014. Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

PEW RESEARCH CENTER

Data shows that (in 2014) 38 percent of democrats see the Republican Party as unfavorable (compared to 16 percent in 1994), with 27 percent seeing the opposing party as a threat to the nation's well-being. Similarly, (in 2014) 43 percent of republicans see the Democrat Party as unfavorable (compared to 17 percent in 1994), and more than a third see the opposing party as a threat to the nation's well-being (Pew Research Centre 2014). Keeping this in mind, it is very difficult to expect that public deliberation in conditions of group polarization and crippled epistemology can yield results of decent epistemic quality.

White and Ypi are aware of these problems, yet their position seems to be immune to these objections since they are not defending the standard account of epistemic democracy. They acknowledge group polarization; yet do not think that it endangers their position. "Though deliberation with epistemic trustees leads to group polarization (*rather than improving the quality of arguments*), whether it is bad depends on the nature and value of one's commitment" (White and Ypi 2016: 96). However, it seems that group polarization (and crippled epistemology that follows), undermines the mechanisms needed for public justification and public reason-giving. Namely, partisans start seeing their political opponents as enemies—as stupid or malevolent people (or both). Consequently, they start seeing public justification as unnecessary—they ask themselves why they would owe public justification to stupid or malevolent people. Group polarization represents a threat not only to the (standard) epistemic conception of deliberative democracy, but also to virtually any form of democracy that relies on public deliberation as a method of political justification, including White's and Ypi's position.

Some might argue that group polarization represents a problem that can easily be settled. Though political parties might prefer to keep their supporters loyal and isolated from the arguments coming from the other side (crippled epistemology), they also want to win elections. And if a party wants to win elections, it will have to embrace and at least incorporate the views of the general, non-polarized population in its political program. The general public will thus keep parties from polarizing (White and Ypi 2011, Biale and Bistagnino 2018). The empirical data from the US, however, points in the opposite direction. Along with political parties, the general public is polarizing as well, with leaners being more similar to partisans than to other independents. Polarized news media and echo chambers created by social media have a strong impact on non-partisan citizens, who start behaving similar to their partisan fellows (Pew Research Centre 2014). It seems that partisanship (as a road to group polarization and crippled epistemology) introduces problems for all forms of deliberative democracy, including both the epistemic view I defend and the public-reasoning view embraced by White and Ypi. Knowledge transmission mechanisms and new epistemic resources partisanship offers are not enough to preserve its epistemic value. If we want to argue that partisanship increases the epistemic quality of democratic decisions (but also if we want to argue that it is compatible with public reason-giving), we have to address group polarization and offer some mechanisms to reduce its damaging effects.

4. *Internal answer*

Building upon White and Ypi's position regarding epistemic value of partisanship, and pushing it even further by adopting the standard account of epistemic democracy, I want to sketch possible remedies for the group polarization problem. This part of the paper examines some party-oriented solutions, while the final part focuses on necessary social and political transformations.

One way of fighting group polarization is by making changes in the internal organization of the party. Changing how party members deliberate and make decisions between themselves might help reduce the damaging effects of group polarization and prevent the development of crippled epistemology. Keeping in mind the devastating effects group polarization can have on the society in general, as well as on the legitimacy and epistemic quality of democratic decision in particular, parties upholding liberal and democratic values should try to prevent epistemically undesirable changes among their members. Changing one's views regardless of reasons (or even despite them) and adopting cognitive and deliberative norms that disregard arguments of others because of *argumentum ad hominem* fallacy represents epistemically undesirable behavior that parties have a duty to prevent, destimulate and sanction. There are a few useful methods borrowed from group

psychology that might help parties fight crippled epistemology. I discuss briefly two such methods in the rest of this part.

First, parties should try to build a critical thinking culture within the organization itself. Instead of simply endorsing and following party's values, opinions and policies, members should be encouraged to deliberate on them and to adopt a wide range of epistemic virtues (e.g. toleration, willingness to argue with others and defend one's views, and disgust towards rhetorical figures that lead to logical fallacies). This recommendation has two goals. First, when widespread deliberation within party is allowed or even encouraged, party members might realize that their organization is not as homogenous as they originally believed—though most (or even all) party members endorse and agree upon certain political values, they might disagree on laws, policies and political decision their party promotes to achieve its aims. Critical thinking culture will thus result in a more heterogeneous organization that is less vulnerable to group polarization and crippled epistemology. Second, members of parties that promote critical thinking will be less stressed if their views are challenged. In fact, members of such parties will often start discussions with their political opponents believing that their position is supported by the best reasons and arguments, while also holding that they have a duty to defend it in front of others.

Second, political parties can use *red teaming* (Sunstein and Hastie 2015), a famous strategy originally used in military training, where red team plays an adversary role and genuinely tries to defeat the primary team in a simulated mission. This is a very useful method for improving the effectiveness of an organization, especially for organizations with strict hierarchy and fixed ways of approaching problems. Many law, computer firms, research institutions (e.g. NASA, IBM, SAIC), as well as government agencies, use similar strategies to understand the weaknesses of their side of a case or a theory. Red teaming is thus an upgraded version of devil's advocate and requires the positions of other parties to be viewed in their full strength (and not as a straw man position intentionally misinterpreted to befame the political opponents), enabling partisans to see that their opponents are neither stupid nor malevolent. However, this method can successfully remove some epistemically damaging effects of crippled epistemology only if red teams are used to challenge the content of parties' political views and not just the rhetorical figures and PR strategies.⁸ Nonetheless,

⁸ We can see members of other political parties as malevolent or stupid and still use red teaming to improve our political performance. Parties' red teams often investigate whether political opponents can use some misinterpretation of political messages during electoral campaigns to turn the very message against the party that issued it. Using rhetorical figures and political propaganda strategies they can strengthen the political message and protect it from misinterpretation, and this can be a good thing from the epistemic standpoint. However, it will not protect us from crippled epistemology. Namely, we can still see our political rivals as malevolent (e.g. as those who try to manipulate the citizens in their favor), and use red teams to

parties have a strong reason to have red teams focusing on the content of their political program. Apart from normative reasons to address the content of opponent's views, parties are aware that the public opinion is shaped not only by their political messages, but also by journalists, political analysts, scientists and other experts who focus on the content of party's political program. Red teams should try to produce the best possible criticism, taking the views of party's political opponents as intelligent and well supported. This should enable party members to review various objections to their political views or to public policies they advance, which is exactly what crippled epistemology normally prevents them from doing. Being aware of your party's weaknesses and of other parties' strengths is a welcomed resource in political competition, and a valuable epistemic state standing in opposition to the norms of crippled epistemology.

This part of the paper indicates some useful epistemic resources that can be employed within political parties to block the effects of group polarization and crippled epistemology. The far-reaching hope is that the use of such resources could decrease the negative epistemic impact of group polarization and crippled epistemology, thus preserving the epistemic value of political parties in a democratic decision-making process. A reasonable worry, however, strikes us when we try to articulate the reason why the parties would be motivated to act in such epistemically virtuous manner. Knowing your weaknesses and other parties' strengths can be useful, but why should most members of a party be aware of this? Why not simply have a small group of experts focusing on possible weaknesses of our proposals and advising the leadership of the party, with most of the party members still polarized and motivated to vote since they see opposing parties as stupid and dangerous? It seems that, in order to block the effects of group polarization and crippled epistemology, we need more than a list of internal measures the party can (but does not have to) employ. We need to address the problems from a wider perspective.

Though many hold that political parties should implement the methods described in the previous part, no one really expects that they will actually do so. Endorsing these methods might be like cooperating in a prisoner's dilemma case—it is an option everyone would benefit from, yet also an option no one will embrace unless there is an instrument that will ensure that everyone embraces it. This is why Sunstein (2007, as cited in Talisse 2017: 113) prescribes the introduction of “legal measures that could limit a doxastic group's capacity to enclave”. He clearly does not count on already polarized groups to welcome the opinions and critiques of their political opponents, but instead aims for an institutional design that would impede the creation of echo chambers within

protect our agenda from simple rhetorical criticism, without engaging the content of our political views or the substantial critiques articulated by our political opponents.

groups.⁹ Following Sunstein's example, one way of fighting group polarization would be to introduce legal measures that ensure the establishment of critical thinking culture within political parties. This is an internal answer to the group polarization problem since, even though the regulation comes from outside the party, it is concerned with how the intra-party deliberation is shaped. I am not inherently against such regulation but I do not think it is up to political philosophers to devise exact legal means that could be used. Furthermore, there is a justified skepticism whether such internal answers will succeed in fighting crippled epistemology in an already polarized world.

5. *External answer*

So far we have focused on institutional and legal measures that could help prevent the development of crippled epistemology within political parties by re-shaping the intra-party deliberation. Alternative solution, one briefly discussed in this part of the paper, is to change the external conditions that lead to crippled epistemology. Namely, crippled epistemology perpetuates itself when citizens attempt to protect their beliefs by "keeping themselves in the company only of others who share their beliefs" (Hardin 2002: 10). This is the reason why extreme groups often try to prohibit or limit their members' interaction with citizens outside the group—they promote extreme religious dogmas, conspiracy theories, pseudo-histories and other epistemic mechanisms to keep their members in the company of only like-minded people. Many closed religious sects, for example, use similar mechanisms to discourage the interaction between their members and the outside world. Such organizations are usually totalitarian in a sense that they try to regulate every aspect of members' private and public life.

Major political parties, on the other hand, usually cannot afford to be as epistemically closed as religious sects. Their members are often also members of various non-political associations, organizations and groups. One might be a member of some political party, but she might also enjoy football and be a member of a local team's supporters' club, sing in a neighborhood's choir, be engaged at the workplace as a member of a trade union, spend her free time talking about science fiction within a book discussion club, and defend animal rights as a member of some animal welfare association. Though these are all non-political organizations, it is very difficult to believe that some political issues will not be (at least partly) discussed or addressed there. Even if there is no substantive political discussion among members of such associations, members will very likely know political views of other members within the organization. They will share some common goals with other members who do not share their political views, they will work together to

⁹ For example, Sunstein holds that websites of extreme parties and groups should be legally required to incorporate links to opposing websites.

achieve them, and they will be able to see that these members, though they are their political opponents, are neither stupid nor malevolent. The same goes in both directions: members of supporters' clubs of opposing football teams might avoid the danger of group polarization by being members of the same political party and collaborating in the local branch of the party.

When citizens are members of various organizations, associations and informal groups, mixed with people holding different political views, political deliberation will not be closed within a single homogenous deliberating group. Even if it remains closed, the danger of crippled epistemology (characterized by demonization of political opponents which leads to the lack of social interaction with those outside the group) will be removed.

We should aim for an institutional design that promotes and protects citizens' plural identities. This calls for laws and public policies that help strengthen and further develop civil society associations, particularly those who have citizens with different and opposing political views among their members. Namely, such associations serve an additional purpose (apart from the one they have identified as their main aim)—they help us impede group polarization and block crippled epistemology.

6. *Conclusion*

Political parties gather citizens with similar political views. This is simultaneously the source of their epistemic worth and the epistemic danger they represent to the quality of democratic decision-making process. On the one hand, having like-minded members facilitates the transmission of knowledge and helps in the development of hermeneutic resilience. On the other hand, political parties are often seen as vessels of polarization since they gather like-minded citizens and often try to present their opponents as stupid or malevolent. Group polarization and crippled epistemology endanger both the epistemic and the moral value of democratic decision-making procedures. They represent a threat to all forms of deliberative democracy, including both epistemic and non-epistemic approaches to the value of public deliberation. Finally, they threaten to overwhelm the positive effects of partisanship, turning political parties into enemies of public deliberation, i.e. into entities that produce more epistemic harm than epistemic good. Introducing certain internal norms (the promotion of critical thinking culture and red teaming within a political party) and external conditions (institutional promotion and protection of citizens' plural identities through special support for civil society organizations) could reduce epistemically and morally damaging features of group polarization and crippled epistemology.

References

- Atchison, T. 2012. "Distrusting Climate Science: A Problem in Practical Epistemology for Citizens." In J. Goodwin (ed.). *Between Scientists and Citizens*. Iowa City: Great Plains Society for the Study of Argumentation: 61–73.
- Biale, E. and Bistagnino, G. 2018. "What Role for Experts in Democratic Politics? A Discussion from the Epistemic Function of Political Parties." [Lecture] *White&Ypi Symposium. Summer School "Equality and Citizenship"*. Rijeka: University of Rijeka.
- Bohman, J. 1998. "The Coming Age of Deliberative Democracy." *Journal of Political Philosophy* 6, (4): 400–425.
- Brady, D. W., Ferejohn, J. and Harbridge, L. 2008. "Polarization and Public Policy: A General Assessment." In P. S. Nivola and D. W. Brady (eds.). *Red and blue nation? Volume One: characteristics and causes of America's polarized politics*. Washington: Brookings Institution.
- Cerovac, I. 2016. "Epistemic Value of Public Deliberation in a Democratic Decision-making Process." *Philosophical Alternatives* 25 (4): 5–16.
- Christiano, T. 2012. "Rational deliberation among experts and citizens." In J. Parkinson (ed.). *Deliberative Systems: Deliberative Democracy at the Large Scale*. Cambridge: Cambridge University Press: 27–51.
- Cohen, J. 1986. "An Epistemic Conception of Democracy." *Ethics* 97 (1): 26–38.
- Elga, A. 2007. "Reflection and Disagreement." *Nous* 41 (3): 478–502.
- Estlund, D. 2008. *Democratic Authority*. New Jersey: Princeton University Press.
- Franken, A. 2004. *Lies and the Lying Liars Who Tell Them: A Fair and Balanced Look at the Right*. New York: Plume.
- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Fricker, M. 2013. "Epistemic justice as a condition of political freedom." *Synthese* 190 (7): 1317–1332.
- Graham, P. 2008. *How to Disagree*. PaulGraham.com. (March 2008) [Online] [Retrieved: November 1st, 2018].
- Hardin, R. 2002. "The crippled epistemology of extremism." In A. Breton (ed.) *Political Extremism and Rationality*. Cambridge: Cambridge University Press.
- Layman, G. C., Carsey, T. M. and Horowitz, J. M. 2006. "Party Polarization in American Politics: Characteristics, Causes, and Consequences." *Annual Review of Political Science* 9: 83–110.
- Marti, J. L. 2006. "The epistemic conception of deliberative democracy defended: Reasons, rightness and equal politicalala." In J. L. Marti and S. Besson (eds.). *Deliberative Democracy and Its Discontents*. Hampshire: Ashgate Publishing: 27–56.
- Nussbaum, M. 2000. *Women and human development: The capabilities approach*. New York: Cambridge University Press.
- Pavićević, Đ. and Simendić, M. 2016. *Disciplinovanje demokratije* [Disciplining Democracy]. Belgrade: Fabrika knjiga.
- Peter, F. 2011. *Democratic Legitimacy*. London: Routledge.
- Pew Research Center. 2014. *Political Polarization in the American Public*. (June 2014) [Online] Available at: <http://www.people-press.org>

- org/2014/06/12/political-polarization-in-the-american-public/ [Accessed: June 2018]
- Prijic-Samaržija, S. 2011. "Trusting experts: Trust, testimony and evidence." *Acta Histriae* 19 (1–2): 249–262.
- Prijic-Samaržija, S. 2017. "The role of experts in a democratic decision-making process." *Etica & Politica / Ethics & Politics* 19 (2): 105–140.
- Rousseau, J. J. 1997. "Of the Social Contract." In V. Gourevitch. *The Social Contract and Other Later Political Writings*. Cambridge: Cambridge University Press: 39–152.
- Sen, A. 2003. *Inequality Reexamined*. New York: Oxford University Press.
- Solomon, M. 2006. "Groupthink versus the wisdom of crowds: The social epistemology of deliberation and dissent." *The Southern Journal of Philosophy* 44 (1): 28–42.
- Sunstein, C. 2002. "The law of group polarization." *The Journal of Political Philosophy* 10 (2): 175–195.
- Sunstein, C. 2005. "Group judgments: Statistical means, deliberation, and information markets." *New York University Law Review* 80: 962–1048.
- Sunstein, C. 2009. *Going to Extremes: How Like Minds Unite and Divide*. Oxford: Oxford University Press.
- Sunstein, C. and Hastie, R. 2015. *Wiser: Getting Beyond Groupthink to Make Groups Smarter*. Boston: Harvard Business Review Press.
- Talisse, R. 2009. *Democracy and Moral Conflict*. New York: Cambridge University Press.
- Talisse, R. 2017. "New trouble for deliberative democracy." *Les ateliers de l'éthique / The Ethics Forum* 12 (1): 107–123.
- White, J. and Ypi, L. 2011. "On partisan political justification." *American Political Science Review* 105 (2): 381–396.
- White, J. and Ypi, L. 2016. *The Meaning of Partisanship*. Oxford: Oxford University Press.

Populists, Samaritans and Cosmopolitans. What is the Right Alliance?

NENAD MIŠČEVIĆ

University of Maribor, Maribor, Slovenia

Central European University, Budapest, Hungary

In the last decade the international situation has been marked on the one hand by refugee crisis, and on the other by right-wing populist reaction to it. This constellation forces a new playground for the traditional philosophical cosmopolitan–nationalist debate. The moral and political issues raised in this new context concern duties to “strangers at our doors”, and these duties and the awareness of them are the first step in a cosmopolitan but realistic direction. Cosmopolitanism now has to start as “samaritan” cosmopolitanism, openness to and engagement for the close and present strangers. Once the present urgent problems are on the way to be solved, we should turn our attention to deeper causes of the crisis. These causes are the evils traditionally discussed by cosmopolitan authors, from dramatic North-South inequalities, to exploitation and warmongering done by the richest countries. The initial samaritan motivation naturally leads to attention to deeper issues, and toward a more ideal cosmopolitan theory. The resulting Samaritan-to-deeper-measures model fits well with Ypi’s engagement with the principle/activism divide, and offers a way of understanding, and hopefully, overcoming it. At the meta-level it connects the appeal to empathy as the relevant moral sentiment and the more rationalist, contractualist justification of global justice.

Keywords: Populism, refugee crisis, Samaritanism, cosmopolitanism.

1. Introduction

The topics of cosmopolitanism and the issues tied to cosmopolitanism in practice, to which this paper is dedicated, are topics on which Lea Ypi has been working for decades, and that are addressed in her original and challenging book on global justice (Ypi 2012), her (co-) edited volume on migration (Fine and Ypi 2016), and in a series of papers. Some issues are also connected to the more general theoretical con-

trasts, like the one between the ideal and non-ideal theory and the like (some addressed in the book presented in Rijeka in 2018 by Jonathan White and Ypi (2016)).¹

What shall we mean by “cosmopolitanism”? In her book, Ypi discusses “[t]he shift in the use of the term ‘cosmopolitanism’ from an associated with the conduct of single individuals to a politically relevant interpretation of justice” (2012: 27). She characterizes cosmopolitanism as focusing upon global distributive equality (2012: 104). This core idea belongs to “political” cosmopolitanism (as described, for instance, in *Stanford Encyclopedia* entry and sources mentioned there). But the tone of Ypi’s book suggests a strong moral motivation, and thus a link with “moral cosmopolitanism”, which I shall assume in the sequel.

Of course, I am aware there are a lot of problems that arise for cosmopolitanism, the ones that has been formulated by its critics. For instance, Miller proposes a dilemma for cosmopolitans:

So cosmopolitanism as a moral outlook seems to be profoundly ambiguous. In its strong form it readily excludes any preference for one’s compatriots, but by simultaneously ruling out other forms of partiality that are integral to a worthwhile human life, it becomes hard to accept. In its weak form, by contrast, it reduces to a broad humanitarianism that does not rule out anything much at all beyond repugnant ideologies that regard some human lives as of no value. The interesting question is whether we can find some intermediate view. (Miller 2016: 24)

Ypi has an answer: Distinguish the level of principles and the level of action, the ideal and the non-ideal. The ideal (and principles) is the right level for cosmopolitanism. But how shall we connect the two? I think I have an answer to suggest that agrees with Ypi’s distinction of levels, but stresses the possibility of passage from one to the other.

In order to introduce my proposal, I shall start from the fact that a new playground has been opened for and made obligatory to cosmopolitan reflection. A few decades ago philosophers were writing about post-communist conflicts going all the way into wars (like the post-Yugoslav one) and international penal justice applying to the war crimes (Hague and Rome above all), and about the cosmopolitan promises of supra-national bodies, like the EU. The hope in a relatively egalitarian liberalism was in the air.

Now, the most urgent problems are different. On the side of human suffering, it is immigrants and refugees in general that are in focus. On the opposite side, the one of rejection and national-cultural egoism one encounters the mass success of populism. Someone might object that the problematics of wars is still alive, and indeed in a dramatic form, say in the Middle East, from Mediterranean to Afghanistan, and that the idea of an allegedly new playground is part of a myopic Western perspective. (A follower of Ingram’s more pessimistic moments, e.g. (2013: 18), might argue for such an objection.) If this is the case, please

¹ Thanks go to Lea Ypi and to the organizers of the conference, Elvio Baccarini and Ivan Cerovac.

relativize the formulation “the new playground” into “the new playground-for-us”, countries where the analytic political-philosophical discussion is going on.

My proposal starts from trans-national, and in this sense cosmopolitan interaction, the activity of helping strangers, refugees, immigrants in general. I shall borrow from J. Waldron the metaphor of the Good Samaritan, who helps the suffering individual not belonging to his tribe:

We tend to think carelessly that the moral functions of the state must be easier to explain on an affinity model. But this is not the case. The most demanding moral requirements are those that insist on our taking care of strangers and doing justice to those with whom we are not already bound by ties of kinship. (Waldron 2011)

The key here to all this—in the rescue cases—is something like proximity, the persons in question being there, on the spot: “[A] certain Samaritan, as he journeyed, came where he was” (Waldron 2000: 1075). This all sounded strange, at least to me in 2000, when I read Waldron; suddenly, with the immigration crisis and the populist explosion, it became most realistic and actual (see also Waldron 2011 and Valentini 2015). The Good Samaritan is the right intermediate figure, connecting the local and the cosmopolitan.

So, I see good samaritanism as a variant of cosmopolitanism (or as its closest ally). More importantly, *it is the variant (or ally) that is crucial in the times of populist explosion. It takes us from the local to the universal, and merges principles and practice in a constructive way.* The road to take is from samaritanism to deeper cosmopolitan measures. So, my framework here is the issue of cosmopolitanism in the populist age, but towards the end of the paper, I shall pass to the more standard variants of cosmopolitanism, where I agree with Ypi a lot.

Here is then the preview. The first sub-section of the next section introduces populism—the relatively new player, focuses upon its right-wing variety and briefly discusses what it is and how it functions, connecting it very briefly to issues of nationalism and communitarianism. The second sub-section turns to a crucial example: Immigration and immigrants, and the populist challenge of the refusal of potential immigrants, focusing on the work of David Miller as the opponent and philosophical guide to the problems. The second section is dedicated to the proposal for a cosmopolitan solution, to be called “the Samaritan-to-deeper-measures model”. It contains two parts, first, the one concerning immediate Samaritan duties, here-and-now, the project of Samaritan cosmopolitanism. The second part turns to deeper causes of the migration disaster, and briefly mentions the standard longer-term cosmopolitan solutions in terms of peace and distributive justice, which brings us back to Ypi and her two level picture. We propose to combine it with a two level view of justification, combining fitting sentiments, like empathy, at the basic level and rational contractualist justification at the higher, more theoretical level. Finally, the conclusion sum-

marizes the issues awaiting those who deal with cosmopolitanism in practice, either through activism, or through reflection of both.

2. *The new playground*

2.1 *Populism*

The present day populism is a relatively new anti-cosmopolitan political machine, a spectre, to quote Gellner and Ionescu: “A spectre is haunting Europe—the spectre of populism.” (Gellner and Ionescu 1969: 1). Leaders like Orban, Trump, Erdogan and their likes are to a large extent shaping the global situation; together with the refugee crisis, populism is determining the new playground that is surrounding us. But, how should we think of it? What is it, in the first place?

The first thing to note about the notion of populism is that it is very thin. It covers all sorts of movements and ideologies suspicious towards elites and friendly to the wide masses of “people”. Margaret Canovan, in her 1981 book *Populism*, has suggested seven different “types” divided into two major categories:

Agrarian Populism

1. Farmers’ radicalism (e.g., the U.S. People’s Party)
2. Peasant movements (e.g., the East European Green Rising)
3. Intellectual agrarian socialism (e.g., the Narodniki in Russia)

Political Populism

4. Populist dictatorship (e.g., Peron)
5. Populist democracy (i.e., calls for referendums and “participation”)
6. Reactionary populism (e.g., George Wallace and his followers)
7. Politicians’ populism (e.g., broad, non-ideological coalition-build- ing that draws on the unificatory appeal of “the people”).

As noted by authors like Cass Mudde:

Populism is understood as a thin-centered ideology that considers society to be ultimately separated into two homogeneous and antagonistic groups, “the pure people” versus “the corrupt elite,” and which argues that politics should be an expression of the *volonté générale* (general will) of the People. (Mudde 2007: 23, see also Mueller 2016)

What this richness and variations in Canovan’s taxonomy, read together with Mudde’s and Mueller’s characterizations, suggest is that populism is probably not a political kind, the concept is too thin, plus vague and general. Right-wing populism *is*, in contrast, a political kind (as is, probably, its left-wing counterpart, which we shall not need to discuss here).² So, let us concentrate upon right-wing populism, the only one relevant for our topic. Right-wing populism is exclusively focused upon one’s community. But the question is, which one? What about nationalism? Is the relevant community the national one, as Taguieff

² But see, for example, the chapter on “Leftist populism” (chapter six) of March (2011), Aslanidis (2017), and Ingram (2017).

(2015) would have it? Interestingly, in the US populism is more like pro-American nationalism or patriotism, while in Europe, it is typically concentrated upon wider belonging, like belonging to Christian civilization in contrast with the Muslim one, and the like.

Let me illustrate. Start from the US populism. Randall Curren (Forthcoming) gives a depressing overview. He notes that the leading populist activist groups, like the Tea Party, are “composed of people who believe that the government is plotting to deprive Americans of their liberties” (2019: 38–39). “Intensely focused on the federal government as its chief enemy, the Patriot movement swelled when the nation was led by a black man suspected of being a foreign-born Muslim and worse” (39). Its militia groups have engaged in armed standoffs with federal authorities and contested federal control of public lands. Its brand of patriotism is focused on the gun rights and the like, he remarks. He talks about Trump’s “pluto-populism” and concludes by noting the following:

This chapter has argued that expressions of patriotism in the USA following the attacks of September 11, 2001, differed in ways that reflected preexisting social, political, and religious divisions. These revolved around the role of Christian fundamentalism and theological variants of it that influenced the Bush administration’s response to 9.11, the role of Southern regionalism and race in shaping US policy and citizenship, and the reactionary movements and economic polarization that set the stage for the emergence of populism in both the USA and Europe following the financial collapse of 2008. The events of 9.11 are identified as a landmark within larger overlapping periods in which reactionary fundamentalist and libertarian movements have emerged, together with the declining economic fortunes of Western societies, waves of immigration, and declining trust in public institutions. (Curren forthcoming: 38–39)

Overgeneralizing from some nationalistic features of some populist movements, authors like Taguieff (2017) conclude that populism is just a manifestation of nationalism, its “vengeance” for phenomena of globalization, as he puts it in the introduction to his book. Benjamin De Cleen in his “Populism and nationalism” quotes earlier analogous proposals. He notes that Stewart (1969: 183) “goes as far as” to call populism “a kind of nationalism”. And he lists other examples (Akkerman 2003: 151, Jansen 2011 and Taguieff 1997: 15). He then explains why this strategy is wrong, why populism is not nationalism:

Populism is a discourse centred around the nodal points ‘the people’ and ‘the elite’, in which the meaning of ‘the people’ and ‘the elite’ is constructed through a down/up antagonism between ‘the people’ as a large powerless group and ‘the elite’ as a small and illegitimately powerful group. Populism is a claim to represent ‘the people’ against a (some) illegitimate ‘elite’, and constructs its political demands as representing the will of ‘the people’ (for similar definitions see Laclau 2005a, 2005b; Stavrakakis 2004, Stavrakakis and Katsambekis 2014). (Rovira Kaltwasser et. al. 2017: 309)

Let me add that one can note that populist parties like German AfD primarily stress the alleged “civilizational” contrasts and then connect

them to more patriotic or nationalist slogans. A typical poster for Federal elections (in this case those that were held in Germany in September 2017) show young women in bikinis, with the simple text: “Burkas? We are for bikinis.” At the bottom, comes the more patriotic encouragement: “Germany, have confidence in yourself!” Other typical posters refer to alcohol drinks: We are not like Muslims, we love brandy, local wine and the like. And again “Germany, have confidence in yourself!”

So, if populism can be wider than nationalism, although connected with it, how should we characterize it? Harald Stelzer has been stressing common elements of communitarianism and populism, and I agree with him. He talks about criticism of modernization processes (involving alleged dissolution of communities, of embedded individual identity in the community and then moral chaos, all this critique imbedded in a criticism of liberal understanding of democracy). He next lists institutional orientation (involving emphasis on participation, demanding solidarity and social order). Finally, there is the assumption of the homogeneity of collectives (with fear of cultural dissolution, cultural particularism and the shared notion of the good).³

This does not make right wing populism into communitarianism, since communitarianism is primarily a philosophical standpoint, not political movement. For populist attitude we need a related but not synonymous term, more tied to politics than to philosophy. Call it “communitarian loyalty” or “strong communitarian loyalty” if you prefer.

So, cosmopolitanism is confronting new problems in the populist age marked by migration crisis and the like. It makes us aware of problems with classical cosmopolitan answer. The ideas of global governance, global economic justice and global justice in general do not speak to the burning issues of the populist age. Cosmopolitanism faces this new playground and the populist challenge. How should we react to it? The crucial example is immigration and immigrants, the favorite topic of population urging.

2.2 *Immigration*

In Europe immigration is probably the main topic of populist uproar. In the US it is one the main topics. So, immigration plus populist reaction are at this moment the main testing ground for cosmopolitan views. Let me start very briefly (with apologies) with anti-cosmopolitan pro-national side, pointing to my disagreements along the way.

Some authors, on the moderate pro-national side, like David Miller (particularly in his 2016), claim that national responsibility to accept immigrant refugees is balanced by considerations of the interest of would-be immigrants and the interests that national communities have in maintaining control over their own composition and character. But in fact, his reservations are much stronger. In discussing outcome

³ Harald Stelzer, “Communitarianism and right wing populism”, a talk on conference in Bled, Slovenia. I thank Harald for sending me his powerpoint.

obligations to the poor, Miller reminds the reader of possible responsibilities of the “people in poor countries who support or acquiesce in regimes that reproduce poverty by siphoning off a large portion of GDP into military expenditure, presidential palaces, and Swiss bank accounts” (Miller Undated: 6).

How seriously should we take this? Take Ghana, mentioned by Miller in the same paper. Its actual GDP per capita is 4.604 US dollars, and it is on the 126th place on the list of countries. Imagine two women from Ghana, one from agricultural, the other from working class family, who migrate with their families to your country. “Sorry ladies, you are co-responsible for corruption of your government. You have acquiesced in a regime that reproduces poverty, that’s the sad fact. So, you merit no help”, you are supposed to argue, from your comfortable seat in the park of Nuffield college. I just cannot believe that Miller would really suggest such a reaction.

He admits that “[...] there may indeed be some refugees to whom redress is owed” but points to the danger “of double bind” (Miller 2016: 176): on the one hand, Western powers are blamed for their intervention, on the other, blamed if they don’t intervene.

In any case, he is convinced of the irrelevance of cosmopolitanism:

In Chapter 2, I discuss cosmopolitanism in general terms as a background to the debate over immigration. Here I simply want to indicate why, even if one is convinced by the general arguments in its favor, it may be less helpful than one might suppose in thinking about the practice of immigration, where this involves not only the question “should borders be open or closed?” but a much wider set of issues about the selection of immigrants, the treatment of refugees, integration policy, and so forth. Thinking about cosmopolitan approaches is, however, a good way of focusing on the question of what we should take as given and what we should regard as amenable to change when discussing immigration. How realistic or idealistic should we be? For example, should we take for granted a world made up of separate states in the first place? Should we assume that global inequalities will be roughly as large as they are now? How else might the current international order be changed?

The argument for swallowing a considerable dose of realism here is simply that the immigration issue would either disappear altogether or at least become much less pressing in a world that was configured quite differently from our own. Suppose there were no separate states, but simply administrative districts accountable to a world government of some sort. There would then be no immigration in the sense in which we understand it. (Miller 2016: 16–17)

Miller very reasonably notes that steps have to be taken to reduce the migrant flows themselves to manageable proportions. He suggests that: “[T]his is partly a matter of working with local authorities in the sending states to clamp down on people-smuggling operations, and to better police their own territorial waters (this is most relevant in the case of states such as Turkey which are themselves safe havens for refugees), and partly a matter of improving living conditions and pro-

viding work opportunities around the camps already established near conflict zones” (2016: 272).

...[T]he refugee issue was likely to prove morally excruciating under certain circumstances, and the European crisis appears to confirm that prophecy. No humanitarian could fail to respond to the plight of drowning boat people, or of land migrants who find themselves blocked by border fences and without basic means of subsistence. They are the hikers in the desert from my second chapter. But equally a co-ordinated response by states to the crisis must consider the longer-term consequences of what is now done—the signals it gives and the incentives it creates for those who might want to move in the future. And where states have developed (justified) policies for different categories of immigrants—refugees, economic migrants, temporary workers, and so forth—these policies should not be torn to shreds because of the current emergency. Citizens and government officials alike have to find a way of compromising between these two imperatives: how to rescue those in need of rescue without turning the border into a free-for-all? (2016: 172)

An urgent question, however, is how this is to be done. Nowadays, if you stop the smugglers, people in the threatened countries will be killed, or raped, or enslaved. So what about alternatives to smuggling? How about us making escape easier by organizing transportation for them?

On the pro-refugee side, Kymlicka has been stressing the advantages of the host state being multi-cultural, offering Canada as his prime example. He is rightly enthusiastic about cultures meeting each other, but also wants to save national solidarity, as against „neoliberal multiculturalism”. He sees it as “a progressive political resource”. So, wishing at the same time to save “immigration and multiculturalism” he tries “to identify the prospects for a multicultural national solidarity” (Kymlicka 2015: 3). Here is his advice:

we need to develop a form of multiculturalism that is tied to an ethic of social membership: that is, a form of multiculturalism that enables immigrants to express their culture and identity as modes of participating and contributing to the national society. A solidarity-promoting multiculturalism would start from the premise that one way to be a proud and loyal Canadian is to be a proud Greek-Canadian or Vietnamese-Canadian, and that the activities of one’s group—be they religious, cultural, recreational, economic or political—are understood as forms of belonging, and of investing in society, not only or primarily in the economic sense, but in a deeper social sense, even (dare I say it?) as a form of nation-building. (Kymlicka 2015: 12)

I find the advice correct and convincing, but unfortunately, the main populism infected (or at least under such threat) European states are far from being multicultural to the extent typical of immigrant countries like Canada. We need a workable cosmopolitan political philosophy of immigration. So, let us go back to the main issue. Miller has been arguing that cosmopolitan arrangement is too distant and far away to point in the direction of actual practice. So, what should (we) cosmopolitans do?

Ypi has one answer: The contrast between ideal and non-ideal. I would, for my own part, like to ask the question whether the here-and-now approach could solve it: the immediacy of sight and mutual recognition is the main point, as Waldron suggests:

I argue that the important moral work in the story of the Good Samaritan is not done by any abstract cosmopolitan universalism—which is very easy to lampoon [...] but by the sheer particularity of the accidental conjunction in time and space of two concrete individuals [...] (Waldron 2011: 16)

So, accept refugees at your doors, organize decent life for them, prove that they are not the threat to ‘us’, and be generous in accepting them. Miller, however, has a warning against starting with samaritanism. He points to a wide range of psychological experiments that suggest that people are not good at helping suffering or threatened others even when the others are in their sight. In the chapter of his *Justice for Earthlings* entitled “Are they my poor: the problem of altruism in a world of strangers” he lists a series of depressing psychological experiments, suggesting that people are bad in samaritan situation. Here is the most ironic of many testimonies recorded:

The experimenters witnessed on several occasions the bizarre spectacle of theology students hurrying to deliver a talk on the Good Samaritan and in the process literally stepping over a man who to all appearances had fallen in the street. Changing the cost of helping, in this case the cost of being a few minutes late to give a talk, transformed the subjects’ willingness to be altruistic. (Miller 2013: 190)

I would think that the psychological material just reinforces the point of the Biblical story of the Good Samaritan. In the story, the victim is ignored by a priest (*hiereus tis*, probably of Jewish faith, on his road to the Temple in Jerusalem) and by a Levite who were passing by, and then helped by our hero. The rate of ignoring is thus over 65%, and the psychologists would probably find it realistic. The psychological point is not that everybody would help, but that more would help in presence than in absence, and that is all. The normative point is that urgency produces the duty to aid, and that the primary bearer of duty are persons present in the situation.

So, why not start with the simplest potentially cosmopolitanism engaging situation, the one of Samaritan help?

3. A solution: Samaritan cosmopolitanism

3.1 The proposal: The Samaritan-to-deeper-measures model

Why bother? Well, because migrants might be dying in front of our eyes. In her book Ypi recounts the tragic story of Adonis Musati, a migrant from Zimbabwe, starved to death in Cape Town while queuing at the offices of South Africa’s home affairs refugee centre (Ypi 2012: 107). We have lots of moral obligations to migrants that derive from the past, from their life circumstances and from their general human dignity. But, obligations come alive most clearly in the actual meeting with the likes of Musati. Waldron, who invented the samaritan approach in the ethics of international relations, rightly talks about “the sheer particularity of the accidental conjunction in time and space of two concrete individuals”:

I argue that the important moral work in the story of the Good Samaritan is not done by any abstract cosmopolitan universalism—which is very easy to lampoon [...] but by the sheer particularity of the accidental conjunction in time and space of two concrete individuals [...].

Who is “we” in the concrete case of Rijeka, the place of our conference? On the one hand, the local community, but from a wider perspective, „we” is best construed starting from our supra-state framework, EU, and then proceeding to particular countries.⁴

Who is primarily responsible for helping? We have been illustrating our proposal with the examples of individuals. But, of course, traditionally the official helper will be the nation-state, as Ypi rightly notes. And here there is a difference mostly multicultural vs. mostly monocultural states (Canada vs. Germany (or Poland), and the former might be more ready helpers, as Kymlicka has been suggesting. The third way is to start above the state, and it is dramatically illustrated by the dilemmas concerning the role of the EU in the refugee drama. But we cannot enter these complications here. We shall just look at the dynamic of interaction.

Here, the immediate help comes first, both normatively and causally; just accept the would-be refugees (indeed, the would-be refugees should be helped in leaving their countries and travelling to us). In longer term, staying should involve opportunity for work and training. Those who wish to stay in countries like Slovenia or Croatia, should be allowed to stay (and there should be a quota for each member-state of the EU, and perhaps wider). Distinguish at least three stages, first, the immediate emergency (starvation, freezing, urgent medical problems) and catering to it, second, settlement and learning (on the host’s and the immigrant newcomer’s side), and third, the stage of (some kind of) citizenship, of relatively stable life in the host country.

The differences between economic migrants and refugees exist, but there is a continuum of cases, and a large space in-between, that should tilt our decisions in favor of the needy. Let me say more about these samaritan stages of the trans-national engagement, even risking some small repetition.

Consider the problems of the first stage. The immediate emergency is assumed psychologically to trigger the samaritan reaction, and nor-

⁴ We might need to take a wider look on samaritanism. It suggests a Proximitist Proposal: The density and the entanglement of interaction in a given location or territory is the crucial first condition. Speaking about a suffering person, Waldron writes: “Never mind ethnicity, community, or traditional categories of neighborliness. The fact that you are there makes you his neighbors” (Waldron 2011: 16). But how large is “there”? Once upon a time, with simple means of transportation it was the geographical openness that counted. To take the example of Hungary and south-Slavonic countries, the Pannonian plain was the relevant “there”; in Central Asia it was the old Transoxiana, in particular Ferghana valley. Switzerland and its history illustrates openness for interaction in spite of intervening mountain ranges. Cultural proximity has been offering opportunities for interaction: smaller Slavonic peoples and Russia, the Muslim world from Morocco to Indonesia and the like.

matively to command it. But can the process start at all? With great numbers of immigrants, and with unprepared host countries (see Žižek's doubts below)? Present day Greece and Italy are offering the spectacular proof of practical possibility. Greece, an economically heavily burdened country, is showing hospitality to something like fifty thousand immigrants (almost a million have passed through the county since 2015). Their life is still difficult, but they do survive, and are getting the necessary minimum. Since 2013, Italy took in over 700,000 migrants! (Turkey has taken more than three million of refugees but has not offered them the minimum as we see it from our more Western perspective.) So, in this case, can implies ought, and the antecedent is fulfilled.

At the second stage, once the migrant lives are not threatened, the ideological issues, re-education etc. steps in. Next, and much later come issues of "ideological disagreement" refugee culture should be accommodated as much as possible (in elements whether threaten elementary human not to rights)

At this second and third stage important changes hopefully take place. On the host's side, the initial empathetic reaction connects the local to the (once) distant strangers and to their society and culture: Had someone been helping Adonis Musati, Zimbabwe would become for him/her "the country from which our Adonis came to us". The host would learn in an empathetic, engaged way what the life is like there, how difficult it is to survive, and so on. (To give a Croatian example, our tradition contains sufficiently many multicultural features, most importantly the centuries long presence of Islam, that might serve as a bridge.)

On the immigrant citizen's side, the welcome and the new way of life might produce positive changes. First, our immigrant, call her Saba, learns to appreciate the host country, say Croatia, and the community of Rijeka which has accepted her. Second, she might, after the experiences of both suffering and welcome, develop a better understanding of compassion. And finally, she starts understanding how her new country fits the larger framework. Simon Keller sees the accomplished perspective as "the perspective of the worldly citizen" (Keller 2013: 250).

Now, this Samaritan obligation can function as a preparation for wider, classically cosmopolitan activity. Waldron would disagree; judging by her reaction in the discussion, Ypi might come closer to agreeing.

Ypi presents the cosmopolitan setting as having to do primarily with distributive justice and equality. However, what she has to say in the book can be linked to the issues of peace (a condition of global distributive justice), to some degree of common governance, and probably to multiculturalism.

Of course, these wider issues linked to cosmopolitanism come in once we turn to causes of migration, at its sources. Take Bush-type war-provoking interventionism and the actual tragic profile of the Middle East. The causes have to do with war, extreme poverty (look at Africa) and bad governance (e.g., in Mexico). What is needed, as we all know, is peace plus more: The decent government, some fairness in the

distribution, the respect of human rights. And only a more cosmopolitan arrangement can guarantee this. We need measures significantly deeper than samaritan hospitality; but fortunately, the later might be a preparation for the former.

Here is Kok-Chor Tanm or how:

[...] to rule out patriotic concern in the real world. What the limited patriotic thesis requires, when applied to the nonideal world in which justice is never fully realized, is that patriots ought also to take their duties of global justice seriously, and that they should be striving actively towards a more just world arrangement, if they want their practice of patriotic favoritism to be legitimate. They may show compatriots special concern, but they must also be sincerely attempting to minimize the background injustices by working towards a more egalitarian world. (Tan 2004: 161)

So, we have two theoretical steps, first, accepting samaritanism and second, agreeing with general cosmopolitan ideology. Let us call this “Samaritan-to-deeper-measures model”. It hopefully offers a tentative answer to a more general question raised by Ypi, the issue of the need of two level playground:

Without cosmopolitanism at the level of principle, statist agency is morally indefensible. Without statism at the level of agency, cosmopolitan principles are politically ineffective and motivationally unsustainable. The avant-garde is crucial to both principles and agency. (Ypi 2012: 179)

But how shall the two come together? Principles are not enough—we cannot do without some cosmopolitan activism. The Samaritan-to-deeper-measures model offers an answer. It starts with host experience with refugees and vice versa, where the participants on each side become familiar with the other side and more sensitive (to what we, philosophers, would describe as cosmopolitan principles). The further cosmopolitan steps might, in the good case, lead to deliberative political process in which the need for deeper measures will lead to more cosmopolitan proposals. (The simple, all-too-simple example is the rising awareness of EU administration that a lot of money should be spent on North-African countries in order to take care of the potential migrants there. The similar process is to be expected in relation to Middle East, once the perspective of peace becomes more realistic.)

So, activism (to turn to Ypi’s favorite topic) starts with refugees, the needy, in the vicinity; participants do the Good Samaritan part within the boundaries of state (or a state-like entity like the EU). The activity prompts cosmopolitan widening and the state deeper measures (search for peace, for more economic equality, for common supra-state governance) will naturally fit into the new activist, more cosmopolitan framework. Let us locate this explicitly into framework proposed by Ypi. She talks about how a

[...] dialectical way of conceptualizing associative political relations combines features of the civil society and family models and clarifies the conditions under which political agency would be effective, and the outcome of political actions would be motivationally sustainable. (Ypi 2012: 133)

These features, according to her view, emphasize popular sovereignty and civic education.

Popular sovereignty, on the one hand, allows cosmopolitan interpretations to enter a deliberative political process, enabling the transformation of political institutions in accordance with their normative requirements. Civic education, on the other hand, complements this process by progressively inserting new normative commitments of cosmopolitan pre-existing cultural, political, and historical practices. Both, I suggest, are indispensable conditions if we want global justice to be more than a cosmopolitan manifesto: popular sovereignty for global egalitarian principles to become politically effective in the first place and civic education for them to be motivationally sustained. (Ypi 2012: 133)

The Samaritan-to-deeper-measures model offers both motivation for a political bite of cosmopolitan practices, and for the relevant civic education. The topic of the latter brings us to our last question:

What is the wider, in particular second-order philosophical framework for the model? My personal preference would be a contractualist one, in the wider sense, above all referring to Rawls and Scanlon. The theory of global justice goes well with it, and my preferred version of justification would be the one of slightly idealized parties in the discussion (see Scanlon 2018). He requires that a proposed basic structure be justified to all those who are asked to accept it and notes “that justification must therefore appeal to the reasons individuals have for accepting such institutions based on how their lives would be affected” (Scanlon 2018: 157). He I talk about “the reasons that individuals have for accepting or objecting to institution” (Scanlon 2018: 157).⁵ This contractualist framework is ideal, to my opinion, for specifying the general cosmopolitan principles. However, the question that is particularly relevant here concerns the first, Samaritan part of the model. Where does it fit into second-order ethical-political theory? I would suggest that we look at the structure of the relevant contractualist justification. Among the reasons individuals have for accepting the proposed institutions are self-centered ones. But we also need other-directed attitudes. If Ivana, the Croat, is to accept the idea that we, Croats, have an obligation to help Saba the Zimbabwean, she might need some empathy-sympathy.

Lacking such a morally fitting sentiment, Ivana might simply refuse to admit the obligation for Croats, in the manner made infamous in the right-wing populist discourse: “I just want to close my doors to Africans, no matter in which situation they are!” It is here that the *samaritan considerations become important for the viability to contractualist justification*: The first steps of contractualist reasoning might appeal to constructive, morally positive sentiments, and the empathy-sympathy is the shining example of such a sentiment.

A purist Kantian Scanlonian might object that empathy is not needed. We are free to appeal to idealization and postulate the idealization

⁵ For a fine application to issues of global justice, see (Gilbert 2012).

according to which Ivana is a moral person.⁶ But this is an unnecessary *ad hoc* measure, given that people are actually empathetic in the relevant situations, and we need as much psychological realism as possible: Do not idealize beyond necessity!

A particular episode of empathy is finally justified at the general level of principle agreed upon in the contractualist procedure of justification in which a proposed basic structure is being made acceptable to all those who are asked to accept it. Pure contractualism specifies for us the final state of any particular piece of moral reasoning—the universal acceptance or something of the sort. It is relatively silent on the first stage(s), but stresses perspective taking and explains how it makes one’s reaction more appropriate. The sentimentalist addition fills the void, and suggests how the later achievements are grounded in the initial ones.

Let me note that a similar dialectics seems to appear more generally in the justification of specific attachments, national(ist), patriotic, purely cultural, or class-focused ones. Attachment is a sentiment, or a deep, standing disposition towards sentiment, and, according to my preferred account of moral-political justification, this sentiment has to survive confrontation with other attitudes in a reasonable, open discussion within the contractualist framework.

More needs to be said about the complications of Samaritan situation. So, let me return to the first-order issue and say more, indeed, in a dialogue with my colleagues, in Rijeka and in Slovenia.

3.2 *Objections and replies*

Let me start with the actual discussion in Rijeka, with warm thanks to all participants. First, Chiara Rauceca: “For the action you are recommending, you need strangers at your doors. But many governments are engaged in preventing refugees, and migrants in general, to get to the doors at all. So, what are we to do in this case?”

If you need a name, call this “the Mexican wall problem” in honor of Trump and his bricklayer creativity.

Reply: We should distinguish two elements in our characterization of the “new playground”, with migrants moving to our countries, the factual and the normative element. To start with the latter, the duty to help remains even if the strangers are not literally at our doors; people on boats travelling towards our port are a case in point. In all the cases alluded to by Rauceca, of strangers “close to our doors”, literally or metaphorically, we have the duty to engage in providing them the possibility to get to our doors. Americans should annihilate the Mexican wall. In the extreme Orbanesque case, where we are legally threatened if we try to help them, we have the duty to oppose the threat, in whatever way we can.

⁶ Ulrike Heuer came close to a general version of such a position in a discussion on the general issue of affect and reason in a Dubrovnik conference; I thank her a lot!

The other element is psychological: People are more prone to help suffering others at their doorstep, than at a distance. Here, Raucea's comment has a serious bite; the psychological ease of engagement and the natural rise of empathy-sympathy is blocked, and this is what Orban and his executioners are counting with.

Zsolt Kapelner has articulated a similar worry: What about intolerant and xenophobic media that systematically keep the locals ignorant about the sufferings of the potential strangers-at-the-doors? What would be samaritan obligations in such a case? It's the "communicational Mexican wall", if you like, to connect it with Raucea's question.

Answer: I agree that ordinary citizens cannot do much here. But the situation is especially relevant for intellectuals: It is a task for us to act as public intellectuals, and write, blog, tweet and the like, about the burning issues covered by silence in the media. Internet is offering possibilities unimaginable two decades ago. And again, can implies ought, and produces a version of samaritan obligation especially demanding for us, intellectuals.

Eletra Repetto has articulated a worry often heard in potential host countries, in particular in Central Europe. Here, the local working class, including an army of jobless young and old people, is poor and needy enough, and the activists have enough work protecting its interests. How should we balance the interests of "our" needy with the interests of would be newcomers? Eletra's question comes close to the "progressive dilemma" as formulated by Kymlicka:

In the postwar period, projects of social justice have often drawn upon ideas of national solidarity, calling upon shared national identities to mobilize support for the welfare state. Several commentators have argued that increasing immigration, and the multiculturalism policies it often gives rise to, weaken this sense of national solidarity. This creates a potential "progressive's dilemma", forcing a choice between solidarity and diversity. (Kymlicka 2015: 3)

Answer: As far as the immediate, first stage of helping is concerned, the survival for immigrants is less costly than the normal life of the home needy. Take as example the medical help. In countries I know well, Croatia, Slovenia and Hungary, the main debate concerns relatively costly medical interventions. In contrast, urgent medical help for refugees often concerns much more simple matters that are much cheaper.

Similarly, in Slovenia we have been massively collecting second-hand, somewhat worn out warm garments, that were precious for the refugees freshly arrived from Middle East. In contrast, in the same country, the radical left is taking as a sign of abject poverty possible cases of local retired elderly people who have to wear second-hand worn out garments. I remember how shocked I was when I discovered how many people are doing it in a country I recently visited.

The dilemma gets much more serious with the issue of jobs and long term prosperity. Here, more sociological and economical research is needed, and I leave the issue open.

4. *Conclusion: Cosmopolitanism in practice*

The migration crisis has brought to the attention of wide public the issues connected with the right of immigrant refugees. What duties, if any, do we, members of third countries, have to such immigrants or would-be immigrants? There is a national responsibility to accept people in dire need, but how far does it go? Second, how should we treat cultural differences that become a central issue once the immigrants, in this case the asylum seekers, settle down?

At one end of the spectrum are authors like Slavoj Žižek (2016), who accept in principle the rights of asylum seekers, but demand from them total cultural integration, almost immediately.

Some authors, like Miller, claim that national responsibility to accept immigrant refugees is balanced by considerations of the interest of would-be immigrants and the interests that national communities have in maintaining control over their own composition and character. In discussing outcome obligations to the poor, Miller reminds the reader of possible responsibilities of the “people in poor countries who support or acquiesce in regimes that reproduce poverty by siphoning off a large portion of GDP into military expenditure, presidential palaces, and Swiss bank accounts”. The remedial responsibility should be focused on the responsibility of citizens of rich countries to ensure fairness in cooperation, and to create an international order that would ensure opportunities to develop.

On the other end of the spectrum we have open multiculturalist option of widely opening the doors, and demanding minimum of integration. Our Samaritan-to-deeper-measures model finds its place at this end of spectrum. So, what would cosmopolitanism look like in actual practice in our world, marked by the rise of right-wing populism and refugee crisis?⁷ Our Samaritan-to-deeper-measures model suggests the following: The moral and political issues raised in this new context concern first and foremost the duties to “strangers at our doors” or at least “close to our doors” (literally or metaphorically), and these duties and the awareness of them are the first step in a cosmopolitan but realistic direction. We have argued that the correct immediate cosmopolitan answer to populist threat is the samaritan one: accept refugees, organize decent life for them, and prove that they are not the threat to “us”.

Thus Cosmopolitanism has to start now as “Samaritan” cosmopolitanism, openness to and engagement for the close and present strangers. This can be done within one’s national state and it connects both the within-state activism and its cosmopolitan counterpart. The contrast is familiar from Ypi’s book, but is not as stark as we might fear. The initial empathetic reaction connects the local to the (once)

⁷ See Nowicka and Rovisco (2009). It is concerned with cosmopolitanism “as a practice which is apparent in things that people do and say to positively engage with ‘the otherness of the other’ and the oneness of the world” and with “cosmopolitanism as a moral ideal” (Nowicka and Rovisco 2009: 2) having to do, among other things, with “the possibility of a more just threat and the refugee crisis”.

distant strangers. Once the present urgent problems are on the way to be solved, we should turn our attention to deeper causes of the crisis. These causes are the evils traditionally discussed by cosmopolitan authors, from dramatic North-South inequalities, to exploitation and warmongering done by the richest countries. The initial Samaritan motivation naturally leads to attention to deeper issues, and toward a more ideal cosmopolitan considerations, both in theory and in practice.

At the second-order level, I have proposed a connection between the initial empathetic sentiment that is rational and fitting in the circumstances of strangers-at-our-doors, and the subsequent process of reasoning, contractualist process of justifying the cosmopolitan global-justice proposal to parties concerned. The Samaritan-to-deeper-measures model and its sentimental-plus-contractualist interpretation fit well with Ypi's engagement with the principle/activism divide, and offers a way of understanding, and hopefully, overcoming it. Its second-order counterpart, going from empathy to mutual justification could offer a definitive understanding of the principles-activism connection, crucial for Ypi's project.

References

- Aslanidis, P. 2017. "Populism and Social Movements." In Rovira Kaltwasser et al. (eds.). 2017: 305–326.
- Canovan, M. 1981. *Populism*. New York and London: Harcourt Brace Jovanovich.
- Curren, R. Forthcoming. "Patriotism, Populism, and Reactionary Politics since 9.11." In M. Sardoc, *Handbook of Patriotism*, Springer.
- De Cleen, B. 2017. "Populism and Nationalism." In C. Rovira Kaltwasser et al. (eds.). 2017.
- Fine, S. and Ypi, L. (eds.). 2016. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press.
- Gellner, E. and Ionescu, G. 1969. "A spectre is haunting Europe – the spectre of populism." In *Populism: Its Meanings and National Characteristics*. London: Macmillan.
- Gilabert, P. 2012. *From Global Poverty to Global Equality A Philosophical Exploration*. Oxford: Oxford University Press.
- Goodin, R. 1988. "What Is So Special about Our Fellow Countrymen?" *Ethics* 98 (4): 663–86.
- Ingram, J. D. 2013. *Radical Cosmopolitics. The Ethics and Politics of Democratic Universalism*. New York: Columbia University Press
- Ingram, J. D. 2017. "Populism and Cosmopolitanism." In Rovira Kaltwasser et al. (eds.). 2017.
- Jansen, R. S. 2011. "Populist mobilization: a new theoretical approach to populism." *Sociological theory* 292: 75–96.
- Keller, S. 2013. "Worldly Citizens: Civic Virtue Without Patriotism." In G. Brock (ed.). *Cosmopolitanism versus Non-Cosmopolitanism: Critiques, Defenses, Reconceptualizations*. Oxford: Oxford University Press: 239–253.
- Laclau, E. 2005a. *On Populist Reason*. London: Verso.
- Laclau, E. 2005b. "Populism: What's in a Name?" In F. Panizza (ed.). *Populism and the Mirror of Democracy*. London: Verso: 32–49.

- March, L. 2011. *Radical Left Parties in Europe*. New York: Routledge.
- Miller, D. 2007. *National Responsibility and Global Justice*. Oxford: Oxford University Press.
- Miller, D. 2013. *Justice for Earthlings: Essays in Political Philosophy*. Cambridge: Cambridge University Press.
- Miller, D. 2016. *Strangers in Our Midst: The Political Philosophy of Immigration*. Cambridge: Harvard University Press.
- Miller, D. Undated. "Who is Responsible?" Accessible at jotamac.typepad.com/jotamacs_weblog/files/who_is_responsible.doc
- Moore, M. 2010. "National Commitments and Universal Duties: On the Interrelationship between Domestic and Global Justice." In B. Keith and S. O'Neill (eds.), *After the nation? Critical Reflections on Nationalism and Postnationalism*, Basingstoke: Palgrave Macmillan.
- Mudde, C. 2007. *Populist Radical Right Parties in Europe*. Cambridge: Cambridge University Press.
- Mueller, J.-W. 2016, *What is populism*. London: Penguin and University of Pennsylvania Press.
- Nowicka, M. and Rovisco, M. (eds.). 2009. *Cosmopolitanism in Practice*. Farnham: Ashgate.
- Rovira Kaltwasser, C. et al. 2017. *Handbook of Populism*. Oxford: Oxford University Press.
- Scanlon, T. M. 2018. *Why Does Inequality Matter?* Oxford: Oxford University Press.
- Stavrakakis, Y. 2004. "Antinomies of Formalism. Laclau's Theory of Populism and the Lessons from Religious Populism in Greece." *Journal of Political Ideologies* 93: 253–267.
- Stavrakakis, Y. and Katsambekis, G. 2014. "Left-wing Populism in the European Periphery: The Case of Syriza." *Journal of Political Ideologies* 192: 119–142.
- Taguieff, P.-A. 1997. "Le populisme et la science politique du mirage conceptuel aux vrais problèmes." *Vingtième Siècle. Revue d'histoire* 56: 4–33.
- Taguieff, P.-A. 2015. *La revanche du nationalisme: Néopopulistes et xénophobes à l'assaut de l'Europe*. Paris: Presses Universitaires De France.
- Valentini, L. 2015. "Social Samaritan Justice: When and Why Needy Fellow Citizens Have a Right to Assistance." *American Political Science Review* 109 (4): 735–749.
- Waldron, J. 2000. "On the Road: Good Samaritans and Compelling Duties." *Santa Clara Law Review* 40 (4): 1053.
- Waldron, J. 2011. "The Principle of Proximity." *School of law – public law & legal theory research paper series working paper* No. 11–08.
- White, J. and Ypi, L. 2016. *The Meaning of Partisanship*. Oxford: Oxford University Press.
- Ypi, L. 2012. *Global Justice and Avant-Garde Political Agency*. Oxford: Oxford University Press.
- Ypi, L. 2013. "Cosmopolitanism without If and without But." In G. Brock (ed.), *Cosmopolitanism versus Non-Cosmopolitanism: Critiques, Defenses, Reconceptualizations*. Oxford: Oxford University Press: 76–89.
- Žižek, S. 2016. *Against the Double Blackmail: Refugees, Terror and Other Troubles with the Neighbours*. London: Allen Lane Penguin Press.

On Three Attempts to Rebut the Evans Argument against Indeterminate Identity

RADIM BĚLOHRAD

Masaryk University, Brno, Czech Republic

The goal of this paper is to assess three arguments that have been proposed to rebut the idea that the notion of indeterminate identity is incoherent. In the first part, the author presents Gareth Evans' argument purporting to show the incoherence of indeterminate identity. Next, the author assesses a rebuttal proposed by E. J. Lowe. Although the rebuttal seems sound, Harold Noonan has shown that its scope is limited. After that, a rebuttal by Peter van Inwagen is analysed. The author compares it with Lowe's and shows that consistent application of the principles van Inwagen uses leads to objects having inconsistent properties. In the final part, it is shown that although the answer proposed by Terence Parsons seems superior to both van Inwagen's and Lowe's, its scope is also limited. As a result, Evans' argument seems to stand unrefuted by these three counterarguments.

Keywords: Metaphysical indeterminacy, indeterminate identity, the Evans argument.

1.

Is indeterminacy solely a feature of language, or also a feature of reality? Several decades ago a similar question was asked about modality. Are necessity and contingency merely properties of sentences, or are they also properties of facts? The prevailing view then was that the world has no modal features. These were thought to be characteristic of our descriptions of the world. But through the work of a number of philosophers the concept of modality *de re* has become a respectable part of metaphysics (see van Inwagen 1990: 283). More recently, a similar question has been asked about indeterminacy, with the default view that there is nothing indeterminate about reality; all indetermi-

nacy inheres in our concepts' having insufficiently delineated meanings and extensions. However, a number of philosophers have recently defended the view that indeterminacy is a feature of the world.¹ There are debates about what exactly this view entails. Some philosophers believe that anyone who accepts that indeterminacy is a feature of the world is committed to the view that sometimes it may be indeterminate whether an object x is identical with an object y . That is, some claim that the proponent of metaphysical indeterminacy is committed to indeterminate identity.

On Parsons' definition of indeterminate identity, it is indeterminate whether x is identical with y if and only if

there is no property that x determinately possesses that y determinately does not possess, and vice versa, but there is at least one property that x determinately possesses such that it is indeterminate whether y possesses it, or vice versa, or at least one property that x determinately does not possess such that it is indeterminate whether y possesses it, or vice versa. (Parsons 2000: 31)

To get a clearer idea of what indeterminate identity might mean in reality, consider the following scenario, which employs the idea of indeterminate diachronic identity:

You own a motorcycle—Cyclone. One day you decide to give it a bit of a face-lift, and disassemble it down to the smallest parts. Due to a lack of time, you leave it disassembled on your garage floor for two years. When you return to your project, you find that a number of the parts have been damaged by rust beyond repair. You decide to invest in new parts and end up replacing about half of the components. In the end, you decide to give the motorcycle a brand-new finish—you paint it black. Since the new motorcycle now looks markedly different from the original one, you decide to conclude the grand renovation by renaming it Hurricane.

Is Hurricane the same motorcycle as Cyclone? There are reasons we can cite in favour of the identity, such as the sameness of half of the original parts, and there are reasons against the identification, such as the complete disassembly and replacement of half of the parts.² As a result, it seems the question has no answer. The identity of Cyclone and Hurricane is, in other words, indeterminate.

To illustrate how the example fits Parsons' definition of indeterminate identity, take Hurricane's property of being black. It is now the case that one of the objects determinately has a property, while it is indeterminate whether the other object has it. Hurricane is determinately black, but it is indeterminate whether Cyclone is black. Or suppose you bought the motorcycle in 2010. It is then determinately true that you bought Cyclone in 2010, but it is indeterminate whether you

¹ See, for instance, Baker (2007), Parsons (2000), van Inwagen (1995) and Williams (2008).

² Although I speak of parts here, the force of the example does not depend on any mereological claims. The point is that since the motorcycles differ in parts, they will clearly differ in properties, which are my main focus here.

bought Hurricane in 2010. If Cyclone and Hurricane do not determinately differ with respect to some properties, and let us now assume they do not, then it is indeterminate whether they are identical.

Proponents of the linguistic account of indeterminacy will say that the indeterminate status of the identity statement in question is completely due to the fact that the expressions in the statement have not been defined precisely. We could, if we wished to, stipulate that the name Cyclone only applies to the original motorcycle as long as the engine, or another part for that matter, has been retained. Or we could say that the name will apply to the object as long as at least half of the parts remain the same. On any of these precisifications it would either be true or false that Cyclone is identical to Hurricane, but since we have not determined the precise meanings of the expressions, the identity question cannot be given a determinate answer.

Proponents of the metaphysical account of indeterminacy will claim, in contrast, that the indeterminacy of the identity claim is due to the fact that the facts in the world do not determine the identity claim either way. They do not deny that vagueness is also a feature of language, but maintain that even if we sharpened all the expressions of our language, some questions about facts could still not be given a determinate answer. In particular, there might still be objects of which it is true to say that it is indeterminate whether they are identical, regardless of the way we describe them.

The notion of metaphysically indeterminate identity has struck many people as suspicious. Notably, Gareth Evans argued in an influential paper that it is downright incoherent (Evans 1978). He showed that if we assume that objects are indeterminately identical, a valid argument can be constructed to show that the objects are, in fact, distinct. Let us look at the argument in greater detail.

- (1) $\nabla(a=b)$
- (2) $\lambda x[\nabla(x=a)]b$
- (3) $\neg\nabla(a=a)$
- (4) $\neg\lambda x[\nabla(x=a)]a$
- (5) $\neg(a=b)$ (Evans 1978: 208)³

Premise (1) is the hypothesis of indeterminate identity and states that it is indeterminate whether a is identical to b . (2) follows by property abstraction from (1) and states that b has the property of being such that it is indeterminate whether it is identical to a . The idea is that if it is true that it is indeterminate whether a is identical to b , then b must possess a certain property, namely, the property of being such that it is indeterminate whether it is identical to a . (3) is a generally accepted truism—it is not the case that it is indeterminate whether a is identi-

³ This paraphrase of the Evans argument differs from the original by using the λ notation to express property abstracts where the original uses circumflexed variables. $\lambda x[\nabla(x=a)]b$ reads 'b has the property of being such that it is indeterminate that it is identical to a'. ' ∇ ' is the indeterminacy operator and reads 'it is indeterminate that'.

cal to a . But then it is not the case that a has the property of being such that it is indeterminate whether it is identical to a , as (4) states. Clearly, then, according to (2) and (4), b and a differ with respect to a property. By the contrapositive of Leibniz's Law, if objects differ in properties, they must be different. From this it follows that (5)—it is not the case that a is identical to b . As a result, any claim as to the indeterminacy of identity of objects leads to the claim that the objects are actually different. The concept of indeterminate identity is incoherent.

The proof has generated extensive discussion. Below, I will focus on answers by three philosophers and assess their merits.

2.

In a brief paper, E. J. Lowe (Lowe 1994) suggests a rebuttal to Evans' argument based on the idea that one cannot legitimately infer from the claim that it is not indeterminate that a is identical to a the claim that a does not have the property of being such that it is indeterminate whether it is identical to a , because this property is not determinately distinct from the property, which a possesses, of being such that it is indeterminate whether it is identical to b . Let us look at the details.

Lowe claims that if, according to (2), b has the property of being such that it is indeterminate whether it is identical to a , then a must have the symmetrical property of being such that it is indeterminate whether it is identical to b . But since these two properties only differ by the permutation of ' a ' and ' b ' they cannot be determinately distinct, for a and b are not determinately distinct. But then the claim made in (4) cannot be true: it cannot be true that a does not have the property of being such that it is indeterminate whether it is identical to a (Lowe 1994: 113–114). It is indeterminate that this property is identical to the property of being such that it is indeterminate whether it is identical to b , and a does have this latter property. So, at best, it must be the case that it is indeterminate that a does not have the property of being such that it is indeterminate whether it is identical to a . But in that case the argument fails to locate a definite difference in the properties of a and b and cannot lead to conclusion (5) by the contrapositive of Leibniz's Law.

This is an ingenious rebuttal of the Evans argument, but it has been shown that it does not cut deep enough.

3.

In his response (Noonan 2003), Harold Noonan claimed that Lowe's rebuttal only works for identity-involving properties, such as the ones in the original statement of Evans' argument. However, he maintains that the argument can be formulated so as to involve other properties and that in such cases the rebuttal is ineffective: '... what Lowe is assuming is that the Evansian pattern of argument against vague identity in the world *essentially requires appeal to properties only ex-*

pressible using the concept of identity. But this is incorrect' (Noonan 2003: 115–116).

I will not paraphrase Noonan's own examples but will instead illustrate his reasoning by applying it to the Cyclone–Hurricane scenario. Consider the predicate 'black'. The predicate is true of Hurricane, so it is not indeterminate whether Hurricane is black. But it is indeterminate whether Cyclone is black. So, Hurricane does not have the property of being such that it is indeterminate whether it is black, but Cyclone does have that property. By the contrapositive of Leibniz's Law, Cyclone and Hurricane must be different. Put formally, where 'P' stands for the predicate 'black', 'a' refers to Cyclone and 'b' refers to Hurricane:

- (1*) $\nabla(a=b)$
- (2*) Pb
- (3*) $\neg\nabla Pb$
- (4*) $\neg\lambda x[\nabla Px]b$
- (5*) ∇Pa
- (6*) $\lambda x[\nabla Px]a$
- (7*) $\neg(a=b)$

It is hard to see how one could utilize Lowe's strategy here. This strategy is based on the claim that *a* and *b* only 'differ' in identity-involving properties that are indeterminately distinct, so *a* and *b* cannot differ determinately. But what properties in this version of the argument could be the candidates for the indeterminately distinct properties? There is only the property of being black, and that property does not make any reference to either Cyclone or Hurricane, so the indeterminacy of their identity cannot do the work it does in the above identity-involving properties. In other words, one could not deny premise (4*), which states that Hurricane does not have the property of being such that it is indeterminate whether it is black, on the grounds that this property is indeterminately identical with some property that Hurricane does have. What property would that be?

I find Noonan's argument quite convincing. That is, if we accept that there is such a property as the property of being such that it is indeterminate whether it is black, Cyclone and Hurricane determinately differ in possession of this property and must, as a result, be different objects.

Let us now turn to van Inwagen's solution, compare it with Lowe's and see whether it constitutes an improvement.

4.

It should be noted at the outset that van Inwagen would not accept the Cyclone–Hurricane example as one of indeterminate identity. Van Inwagen's ontology only contains two kinds of objects—simples and organisms (see van Inwagen 1995). Motorcycles are mere simples ar-

ranged motorcyclewise, and, thus, the question of ‘their’ identity never arises. But van Inwagen’s strategy for dealing with Evans’ argument is quite independent of the reasons for his ontological asceticism, so we may ignore this detail.

The central idea of van Inwagen’s response to Evans’ argument is that if objects are indeterminately identical and one of them possesses a certain property, then the other must possess that property at least indeterminately (see van Inwagen 1995: 253). But if that is the case, we will not be able to find a determinate difference between them and, thus, reach the conclusion of Evans’ argument. Let us now look at the details.

Van Inwagen develops a semantics for the language of first-order logic including identity, property abstraction and the sentence operator ‘indef’ which abbreviates ‘it is neither definitely true nor definitely false that’ (van Inwagen 1995: 246). His aim is to show that the use of property abstraction in Evans’ argument can reasonably be considered invalid (van Inwagen 1995: 246). The semantics utilizes three truth values: 1, 0, and $\frac{1}{2}$. Van Inwagen shows that a model can be found in which the step from premise (3) to premise (4) does not preserve definite truth, because it leads from a claim with truth value 1 to a claim with truth value $\frac{1}{2}$, thus making the argument invalid.

It will not be necessary to reconstruct the whole semantical fragment that van Inwagen develops. It is only important to realize why the inference from (3) to (4) is invalid. Let us, first, restate the inference slightly more succinctly. In this section, I will use van Inwagen’s operator ‘indef’ instead of ‘ ∇ ’.⁴

The crucial inference can now be restated as follows:

$$\neg\text{indef}(a=a) \square \neg\lambda x[\text{indef}(a=x)]a$$

I will now paraphrase those components of the semantics that will enable us to evaluate this inference.

1. A *universe* **U** is a non-empty set of objects.
2. A *pairing* on a universe is a (possibly empty) set of two-membered sets (pairs) of members of that universe.
3. Objects are *paired* iff it is indefinite whether they are identical.
4. If a constant ‘*a*’ refers to object A and if A is paired with B, then B is the *fringe referent* of ‘*a*’.
5. The *extension* of an identity predicate contains just the referent

⁴ While writing this paper, I had to make a decision about what notation and terminology I would use to speak about the various theories, because they differ in these respects. Van Inwagen uses ‘indefinitely’ where others use ‘indeterminately’. I generally use ‘indeterminately’ throughout this text, and only when I directly refer to van Inwagen’s definitions and notation do I respect his term ‘indefinitely’ and the operator ‘indef’. Also, like Evans, van Inwagen uses circumflexed variables where others use the λ notation to express property abstracts. I adopt the latter alternative. Finally, van Inwagen deviates from the original statement of the Evans argument by permuting ‘*a*’ and ‘*b*’. I do not adopt this strategy and paraphrase Inwagen’s rebuttal to fit the original statement of the Evans argument.

- of its term; the *frontier* of an identity predicate contains just the fringe referents of its term.
6. The result of prefixing '¬' to a predicate having extension e and frontier f is a predicate having extension $\mathbf{U}-(e \cup f)$ and frontier f .
 7. The result of prefixing 'indef' to a predicate having frontier f is a predicate having extension f and an empty frontier.
 8. The extension and frontier of an *abstract* are the extension and frontier of the predicate on which it is formed.
 9. An identity sentence is:
 - a. true iff something is the referent of both terms;
 - b. $\frac{1}{2}$ iff nothing is the referent of both its terms and the referents of its terms are paired.
 10. An ascription sentence (that is, a sentence in which a property is ascribed to an object) is
 - a. true iff the referent of its subject belongs to the extension of its abstract;
 - b. $\frac{1}{2}$ iff the referent of its subject does not belong to the extension of its abstract, and either (a) the referent of its subject belongs to the frontier of its abstract, or (b) a fringe referent of its subject belongs either to the extension or to the frontier of its abstract;
 - c. false iff neither the referent nor a fringe referent of its subject belongs either to the extension or to the frontier of its abstract.
 11. If 'ϕ' is true, then 'indef ϕ' is false; if 'ϕ' is $\frac{1}{2}$, then 'indef ϕ' is true; if 'ϕ' is false, then 'indef ϕ' is false.
 12. If 'ϕ' is true, then '¬ϕ' is false; if 'ϕ' is $\frac{1}{2}$, then '¬ϕ' is $\frac{1}{2}$; if 'ϕ' is false, then '¬ϕ' is true.
 13. A valid inference form is truth-preserving and does not lead from the value $\frac{1}{2}$ to false. (van Inwagen 1995: 249–251)

Van Inwagen then considers the following model:

{A, B}, {{A, B}}, 'a' ref A, 'b' ref B

On this model, the universe contains only two objects, A and B; these objects are paired, that is, indefinitely identical; and A is the referent of 'a' and B is the referent of 'b'.

The left-hand side of the inference consists of the negation of an indeterminate identity sentence. The embedded sentence '(a=a)' meets condition 9, because the referent of both terms is A, and, therefore, is true. According to 11, 'indef (a=a)' is false. According to 12, '¬indef (a=a)' is true. This is just what we would intuitively expect: it is not true that the identity of a to a is indefinite.

The right-hand side is the negation of a sentence which ascribes the property of *being such that it is indeterminate whether something is identical to a* to the referent of 'a'. Again, intuitively and in accordance with Evans' reasoning we would judge that this sentence is true: object

A does not have the property of being such that it is indeterminate whether it is identical to A. But on the above model the value of this sentence comes out as $\frac{1}{2}$. Let us look at the individual evaluation steps.

At the core of the sentence is an identity predicate '(a=x)'. According to 5, the extension of the predicate contains the referent of *a*, that is, the object A, and the frontier contains the only fringe referent of *a*, that is, object B. Next, according to 7, the extension of the predicate 'indef (a=x)' is the same as the frontier of the predicate '(a=x)', that is, object B, and its frontier is empty. The extension and frontier of the abstract ' $\lambda x[\text{indef}(a=x)]$ ' are the same as those of the predicate 'indef (a=x)', according to 8. Importantly, according to 10, the value of the ascription sentence ' $\lambda x[\text{indef}(a=x)]a$ ' will be $\frac{1}{2}$: the referent of the subject *a*, that is, object A, does not belong to the extension of ' $\lambda x[\text{indef}(a=x)]$ '—we have just seen that that extension contains object B. But a fringe referent of *a*, that is, object B, belongs to the extension of the abstract. According to 10b, this gives the ascription sentence the value of $\frac{1}{2}$. Finally, the negation of a sentence with the value of $\frac{1}{2}$ has, according to 12, the value of $\frac{1}{2}$. Thus, the value of the complete sentence on the right-hand side of the inference is $\frac{1}{2}$. The inference step from premise (3) of the Evans argument to premise (4) does not preserve truth, because it leads from a true sentence to a sentence with the value of $\frac{1}{2}$. That means, according to 13, that the inference is invalid.

The above reasoning can be put less formally as follows. Suppose that we are not dealing with an object indefinitely identical to another object, but with a regular object the identity of which is definite. Call the object C. In such a case, it is quite obvious that C is definitely identical to C, that it has the property of being definitely identical to C, and, as a result, that it does not have the property of being indefinitely identical to C. But here we are not dealing with such regular objects. We are dealing with objects A and B which are indefinitely identical. And the key intuition is that if one of them has a certain property, the other one must 'sort of' have it too' (van Inwagen 1995: 255).⁵ So if B has the property of being such that it is indeterminate whether it is identical to A, then because it is indeterminate whether B=A, A must also 'sort of' have the property of being such that it is indeterminate whether it is identical to A. And this 'sort of' status is formally expressed by the fact that the sentence which ascribes to A the property of being such that it is indeterminate whether it is identical to A has the value of $\frac{1}{2}$ and its negation as well.

Van Inwagen illustrates this reasoning by examples involving empirical properties. In one of them, he describes *the Cabinet*, an infernal philosophical engine which can disrupt the life of anyone who enters in such a way that it is indefinite whether the person who later emerges from it is the same person as the person who entered. The person who

⁵ This does not preclude the possibility that A and B fully share some of their properties.

entered is called Alpha, the person who emerged is called Omega. Since we do not know whether Alpha survived the changes in the Cabinet, it is indefinite whether Omega is identical to Alpha (van Inwagen 1995: 241–242). Suppose further that Omega is hanged when he emerges from the Cabinet. Van Inwagen comments: ‘it is quite definitely true of Omega that *he* dies by hanging. Could it be definitely false of Alpha that *he* dies by hanging? It is hard to see how this could be, given that it is not definitely false that Alpha is numerically distinct from Omega’ (van Inwagen 1995: 253).

Or consider again our Cyclone–Hurricane example. You have painted the motorcycle black. As a result, it is definitely true of Hurricane now that it is black. Could it be definitely false of Cyclone that it is black? Again, it is hard to see how it could, given that it is not definitely false that Cyclone is numerically distinct from Hurricane.

Van Inwagen then concludes: ‘Should matters be different if [the property abstract] contains the symbols “=” and “indef”? I do not see why they should’ (van Inwagen 1995: 254). In other words, even if the property in question is, say, *the property of being such that it is indefinite whether it is identical to a*, the above reasoning still holds. If B has this property, then A must have it indefinitely, because it is indefinite whether it is identical to B. As a result, it is indefinite whether A has the property of being indefinitely identical to A.

In what follows, I will first compare van Inwagen’s strategy with Lowe’s, and then express concerns about its effectiveness. I will show that the consistent application of one of van Inwagen’s key principles leads to objects having inconsistent properties.

5.

There is an interesting parallel between Lowe’s and van Inwagen’s approaches. Both of them attack the inference from (3) to (4). Both argue that from the fact that it is not indeterminate whether *a* is identical to *a*, one cannot infer that *a* does not have the property of being such that it is indeterminate whether it is identical to *a*. But they do it for slightly different reasons. Lowe bases his strategy on considerations related to the identity of properties, which is something van Inwagen does not consider. For Lowe, the property of being such that it is indeterminate whether it is identical to *a* is not determinately distinct from the property of being such that it is indeterminate whether it is identical to *b*. But since object A has this latter property, it cannot be claimed that it determinately does not have the former property. At best it can be claimed that it is not determinate that it does not possess it. But, to repeat, the reason inheres in the fact that the two properties are not determinately distinct.

Van Inwagen seems to suppose that we can reach the same conclusion from the mere fact that the objects that allegedly have those properties are not determinately distinct. He states that ‘if a constant *k*

definitely denotes something x , and there is a y such that it is indefinite whether $x=y$, and y definitely has the property denoted by the abstract F , then ' k has F ' should receive a value of at least $\frac{1}{2}$ ' (van Inwagen 1995: 254). As a result, if b definitely has the property of being such that it is indefinite whether it is identical to a , then since it is indefinite whether $b=a$, the sentence ' a has the property of being such that it is indefinite whether it is identical to a ' must receive a value of at least $\frac{1}{2}$. That is why it cannot be true that a does *not* have the property of being such that it is indefinite whether it is identical to a .

There is one seeming advantage to van Inwagen's approach. We have seen that Lowe's reasoning fails if we formulate the Evans argument using properties not involving identity, for then he loses ground for his claim that the properties in the original Evans argument are indeterminate identical, because they only differ by the permutation of their constants. Nothing in van Inwagen's approach suggests, however, that his reasoning would be limited to identity-involving properties. After all, the above principle applies to properties generally. This opens the possibility of refuting even those versions of the Evans argument that involve regular properties, such as being black. I will attempt to show, however, that there are unwelcome consequences.

6.

Consider again the fact that Hurricane is black. There is nothing indeterminate about this fact. It is just there, standing in front of you, black. Now take van Inwagen's principle that if it is indeterminate whether a is identical to b and a determinately possesses a certain property, then b must possess that property at least indeterminate. As a result, Cyclone has (at least) indeterminate the property of being black. That makes sense, because if the motorbike standing in front of me is determinately Hurricane and indeterminate Cyclone, and if Hurricane is determinately black, then Cyclone must be indeterminate black.

But notice that van Inwagen's principle is a general one and nothing prevents us from using it to reason back from Cyclone to Hurricane. We have now concluded that Cyclone is indeterminate black, that is, that Cyclone has the property of being such that it is indeterminate whether it is black. But Cyclone is indeterminate identical to Hurricane and, according to van Inwagen's principle, any property it has will be such that Hurricane will have it at least indeterminate. So if Cyclone is indeterminate black, it must be true of Hurricane that it is indeterminate that it is indeterminate black. But this is clearly inconsistent with the fact that Hurricane is, right there in front of me, black. How could an object determinately be black and at the same time be such that it is indeterminate whether it is indeterminate black?

To see the problem even more clearly, consider van Inwagen's own example with the Cabinet. Alpha enters the Cabinet. The Cabinet causes changes to Alpha resulting in its being indeterminate whether

Alpha has survived. Then someone, Omega, emerges from the Cabinet and we immediately hang him. Omega is clearly dead. This is as determinate as anything can be. But it is indeterminate whether Omega is Alpha, so, by van Inwagen's principle, it is indeterminate whether Alpha is dead. Good. But if it is indeterminate whether Alpha is dead, by the same principle it must be the case that it is indeterminate that it is indeterminate that Omega is dead. And I do not see how that could be if it is quite clear that Omega is dead.

Or consider that you behead Omega. There he lies, his head separated from his body. Obviously, he is dead and determinately so. Suppose someone doubts: 'Well, I am not saying that it is *indeterminate* that Omega is dead; that is clearly not the case. But what I am saying is that the indeterminacy of Omega's death is indeterminate.' What would that mean? That would mean that *perhaps* it is indeterminate that Omega is dead, but *perhaps* it is not indeterminate, and, as a result, Omega is either determinately dead or determinately alive. I think that, looking at the head lying 3 feet away from the body, these speculations cannot be taken seriously.

To drive the point home, consider an analogy with epistemic certainty. If something is determinately the case, it is certain that it is the case. If something is indeterminately the case, it is uncertain whether it is the case. Suppose we say that it is certain that Omega is dead and then add that it is uncertain that it is uncertain that Omega is dead. That is a paradox. Of course, it would be a more blatant paradox if we only added that it is uncertain that Omega is dead. But even if we weaken the uncertainty of Omega's death by declaring even that fact uncertain, there still remains a grain of uncertainty, which is inconsistent with what we see outside the Cabinet. Uncertainty about uncertainty does not make a certainty.

The advantage of considering the argument with empirical properties is that it shows us clearly that there is an inconsistency in van Inwagen's rebuttal, something which is less clear when we ponder over the cases with identity-involving properties. In the identity-involving cases we rely on our intuitive a priori judgments about the concept of identity. In the empirical-property-involving cases we rely on the evidence of our senses. If someone wants to claim that it is indeterminate that Omega is indeterminately dead, I just point to the corpse. But, ultimately, the fate of these cases must be the same. You are looking at Hurricane. It is there and it is definitely identical with Hurricane. How could it at the same time be indeterminate that it is indeterminately identical with Hurricane?

7.

Could my reasoning be blocked? Could I have taken an illegitimate step? The reasoning consists in two steps. First, I have reasoned from an object definitely having a property to an object (indefinitely identi-

cal to the former object) having the property indeterminately. Second, I have reasoned back from this latter object's having the property indeterminately to the former object indeterminately having this property indeterminately.

Perhaps I have not paid careful attention to van Inwagen's formulation of the key principle. He states that 'if a constant k definitely denotes something x , and there is a y such that it is indefinite whether $x=y$, and y definitely has the property denoted by the abstract F , then ' k has F ' should receive a value of at least $\frac{1}{2}$ ' (van Inwagen 1995: 254). The relevant condition is ' y definitely has the property denoted by the abstract F '. The principle says nothing about the situation when it is indefinite whether an object has the property denoted by the abstract F . So perhaps we are not allowed to reason back from Cyclone to Hurricane and from Alpha to Omega, because neither Cyclone nor Alpha has the relevant property definitely. It is indefinite whether they have them, and the principle does not warrant reasoning back to Hurricane's or Alpha's having that relevant property with a further degree of indefiniteness.

But this sort of reply would seem ad hoc to me. By accepting the validity of the inference from (1) to (2), van Inwagen accepts that the fact that it is indefinite whether a is identical to b (a claim about the indefiniteness of a certain state of affairs) entails that b definitely has a certain property, namely that of being such that it is indefinite whether it is identical to a (a claim about an object having a certain property). Similarly, I do not see why we could not claim that, since it is indefinite whether Cyclone has the property of being black (a claim about the indefiniteness of a certain state of affairs), then Cyclone definitely has a property of being such that it is indefinite whether it has the property of being black (a claim about Cyclone's property). Blocking this inference while retaining the inference from (1) to (2) seems unjustified. But if we accept the inference, then we have to accept that Cyclone definitely has the above property, and we may apply van Inwagen's principle again to reason back to Hurricane having that property indefinitely. Finally, we reach the property of being such that it is indefinite whether it is indefinite whether it is black, which, I claim, is inconsistent with the property of being black.

I conclude that iterating the application of van Inwagen's principle leads to the objects' having inconsistent properties. As a result, van Inwagen's rebuttal of the Evans argument must be abandoned.

8.

We have seen that to avoid ending up with objects with inconsistent properties, the defender of van Inwagen's strategy would have to block iterating the application of the crucial principle. That would mean allowing that Cyclone has the property of being such that it is indeterminate whether it is black, but denying that we could reason back to Hurricane having the property of being such that it is indeterminate

whether it is indeterminate whether it is black. I have argued this restriction is unmotivated.

But a more sweeping strategy has been defended in the literature, which aims to strike the Evans argument at two points. The strategy is to show that the use of property abstracts throughout the argument is illegitimate. This amounts to showing that both the step from (1) to (2) and the step from (3) to (4) are invalid, because it is not legitimate to infer anything about an object's properties from the indeterminacy of its identity. Such a solution has been defended by Terence Parsons (2000) and we will look at it now.

What the Evans argument assumes is that premise (1) reports a fact about *b*, namely the fact that it is indeterminate whether it is identical to *a*, and that we may express the fact explicitly in (2) by ascribing to *b* a property, namely the property of being such that it is indeterminate whether it is identical to *a* (Evans 1978: 208).

But this is not the only way one can look at the situation. The sentence expressed in (1) could be true even if there were not a particular property that *b* has. (1) states that it is indeterminate whether *a* is identical to *b*. The reason this is true is not because *b* has the property of being such that it is indeterminately identical to *a*, and *a* has the property of being such that it is indeterminately identical to *b*. Rather, (1) is true by virtue of it not being determinate whether the properties that one of the objects determinately possesses are determinately possessed by the other, and the properties one of the objects determinately does not possess are determinately not possessed by the other. In other words, (1) is made true by the complex fact or state of affairs in which *a* and *b* occur, and it is not necessary to postulate further properties of the objects to make (1) true. The fact that it is indeterminate whether *a* and *b* are identical is, in short, *fully reducible* to the first-order properties of *a* and *b* and the ways in which *a* and *b* exemplify them, and no *further properties* need be postulated to explain the fact.

Parsons' rebuttal is based on the denial of the principle that for every predicate there is a property the predicate expresses. This is a controversial point, but Parsons argues that it is reasonable to abandon the principle in indeterminacy contexts. Before I elaborate, I should mention that Noonan, who endorses the Evans argument, is well aware of an answer along these lines, but challenges it. He admits that in some contexts, such as intensional contexts, predicates do not necessarily express properties. He considers the predicate 'John believes *x* to be identical with Tully': '... if John believes Tully to be identical with Tully, but does not believe Cicero to be identical with Tully, it does not follow that Tully and Cicero differ in their properties' (Noonan 2003: 144). The reason for this is that the difference only lies in the different ways that the objects are *represented*. But, Noonan claims, friends of indeterminate identity want to say that the identity between *a* and *b* is indeterminate due to how *a* and *b* are *in fact*, not just due to the way

we may represent them (Noonan 2003: 144). In other words, it is easy to understand why the fact that someone believes something about an object does not constitute the object's having a property. But it is not so easy to understand how the fact that object b is indeterminately identical to object a does not constitute a property of b when at the same time it is assumed that the indeterminacy is *de re* and not just in the way we represent the objects.

But Parsons does offer a reason that is independent of the above considerations. The reason can be most clearly seen when we unpack the concept of identity by means of equivalence of property exemplification (Parsons 2000: 50). The sentence

$$(a=b)$$

can be expressed as

$$(P) (Pa \leftrightarrow Pb),$$

that is, as the claim that a and b have exactly the same properties.

$$\nabla(a=b)$$

then renders

$$\nabla(P) (Pa \leftrightarrow Pb).$$

The question of whether there is a property of 'being such that it is indeterminate whether it is identical to a ' then turns into the question of whether the following abstract represents a property:

$$\lambda x[\nabla(P) (Pa \leftrightarrow Px)]$$

The abstract reads 'the property of being such that it is indeterminate whether it has exactly the same properties as a '. The problem with this abstract is, according to Parsons, that it 'must "quantify into" an indeterminacy connective; the abstraction operator on the outside must bind a variable within the scope of the indeterminacy connective' (Parsons 2000: 50). This leads to a problem 'closely associated with the paradoxes of naïve set theory' (Parsons 2000: 50).

In other words, the phrase 'exactly the same properties as a ' actually refers to *all* properties of b including the property expressed by the complete abstract: '... the abstract stands for a property that is in the range of its own property variable' (Parsons 2000: 51).

According to Parsons, this self-referential aspect of the abstract is a good reason why we should be sceptical whether it actually expresses a property that enters the definition of identity.

This is an ingenious response to the Evans argument. But it seems clear to me that, ultimately, its fate is the same as that of Lowe's argument. Perhaps Parsons is right that the predicate 'being such that it is indeterminate whether it is identical to a ' does not express a property, but we have seen that the Evans argument can be formulated without any reference to identity-involving properties. And it is also clear that Parsons cannot make use of his argument against the existence of identity-involving indeterminate properties to argue about other properties

that may figure in the Evans argument.

Let us return to the argument we formulated in part 3, where ‘P’ stands for the predicate ‘black’, ‘a’ refers to Cyclone and ‘b’ refers to Hurricane:

- (1*) $\forall(a=b)$
- (2*) Pb
- (3*) $\neg \forall Pb$
- (4*) $\neg \lambda x[\forall Px]b$
- (5*) $\forall Pa$
- (6*) $\lambda x[\forall Px]a$
- (7*) $\neg(a=b)$

Adopting Parsons’ strategy, we might want to question whether the steps from (3*) to (4*) and from (5*) to (6*) are legitimate. We might object that there is no guarantee that the predicates expressed in (3*) and (5*) express genuine properties, as suggested by the property abstracts in (4*) and (6*). But how could we justify the claim now? In the original formulation of the argument we could unpack the property of identity in terms of the equivalence of property exemplification and show that the property of being such that it is indeterminate whether it is identical to a falls among the properties that the property actually quantifies over. But in this formulation, there is no identity-involving property which could be unpacked. There is the property of being such that it is indeterminate whether it is black and the problem of ‘quantifying into’ does not occur. Since that was the primary reason why Parsons refused to accept the existence of the problematic properties and that reason does not apply here, we are left wondering why the inference should be illegitimate.

Conclusion

The Evans argument is based on the idea that the fact that it is indeterminate whether a and b are identical projects into their properties. While b has the property of being such that it is indeterminate whether it is identical to a , a lacks this property. Evans concludes that this makes a and b distinct. I have looked at three attempts to block this conclusion. Lowe’s response was based on the claim that the identity-involving properties that the Evans argument employs do not make a and b determinately distinct. But, as Noonan has, in my opinion conclusively, shown, Lowe’s rebuttal is toothless when confronted with formulations of the Evans argument which do not employ identity-involving properties. Van Inwagen’s response is based on the idea that indeterminately distinct objects simply cannot differ determinately in the properties they exemplify. So even when b exemplifies the property of being such that it is indeterminate whether it is identical to a , we must conclude that a has this property at least indeterminately, so it cannot be determinately distinct from b . By using examples that, again, avoid refer-

ence to identity, I have shown that this strategy leads to objects having inconsistent properties. Finally, Parsons' rebuttal is based on the idea that the fact of indeterminate identity between *a* and *b* need not mean that they exemplify some further identity-involving properties and the idea that the property of being such that it is indeterminate whether it is identical to *a* is problematic due to self-reference. Even though this is an ingenious response, it is, again, toothless against formulations of the Evans argument that do not employ identity-involving properties, as I attempted to show in the final part of this paper. Consequently, the three arguments I have considered do not threaten the Evans argument and, as a result, its central idea that the notion of indeterminate identity is incoherent, seems to stand unrefuted. None of the three attempts to rebut the Evans argument have shown conclusively that the relation of indeterminate identity has any instances.⁶

References

- Baker, L. R. 2007. *The Metaphysics of Everyday Life: Essays in Practical Realism*. Cambridge: Cambridge University Press.
- Evans, G. 1978. "Can There Be Vague Objects?" *Analysis* 38 (4): 208.
- Lowe, E. J. 1994. "Vague Identity and Quantum Indeterminacy." *Analysis* 54 (2): 110–114.
- Noonan, H. 2003. *Personal Identity*. London and New York: Routledge.
- Parsons, T. 2000. *Vague Identity: Metaphysics and Semantics*. Oxford: Oxford University Press.
- Van Inwagen, P. 1995. *Material Beings*. Ithaca: Cornell University Press.
- Williams, R. 2008. "Ontic Vagueness and Metaphysical Indeterminacy." *Philosophy Compass* 3 (4): 763–788.

⁶ I would like to thank the Czech Science Foundation (GAČR) for supporting this paper with grant no. 17-12551S.

Content, Mental Representation and Intentionality: Challenging the Revolutionary Character of Radical Enactivism

PIERRE STEINER

University of Technology of Compiègne - Sorbonne Université, France

Criticisms and rejections of representationalism are increasingly popular in 4E cognitive science, and especially in radical enactivism. But by overfocusing our attention on the debate between radical enactivism and classical representationalism, we might miss the woods for the trees, in at least two respects: first, by neglecting the relevance of other theoretical alternatives about representationalism in cognitive science; and second by not seeing how much REC and classical representationalism are in agreement concerning basic and problematic issues dealing with mental content and intentionality. In order to expand and exemplify these ideas, this paper presents two heterodox positions on intentionality and on the relations between content and representation. Special attention is paid to the way REC is rejecting these positions: I argue that this rejection reveals common assumptions with classical representationalism, but also undermines the coherence of REC's conception of intentionality.

Keywords: Content, intentionality, mental representation, radical enactivism, 4E cognition.

Introduction

4E—embodied, embedded, extended and enactive—approaches to cognition are fashionable these days (Newen, De Bruin and Gallagher 2018). As their names suggest, these approaches insist on the constitutive importance for cognition of its embodied, embedded, enactive and extended dimensions. By *embodiement*, one means the bodily and organismic realization of cognition, much broader than its neural basis. By *embeddedness*, one means the fact cognitive processes are situated

in a biological and cultural environment that supports and constrains the way they are acquired and exercised. By *enactive*, one points to the fact perception and action are not separated processes: what one does and can do influences the way one perceptually experiences the world.¹ Perceiving is a way to act (Noë 2004). By *extended*, one argues that there are some circumstances in which cognitive processes may be literally realized in some of the environmental structures embodied agents are reliably coupled with. All in all, *cognition is not (only) in the head; it is not (only) a matter of mental representations; it unfolds (or is “enacted”) in the coupling relations or interactions between embodied and living organisms and their social, cultural, linguistic, technological and biological environments* (Clark 1997; Chemero 2009; Hutto and Myin 2013). Notably inspired by original works in robotics, linguistics, ecological psychology, and anthropology, but also by the rediscovery of phenomenology and pragmatism, 4E approaches criticize the internalist, representationalist, individualist, formalist and reductionist tendencies of the dominant paradigm in cognitive science (computationalism).

The theoretical unity of 4E cognition should not be taken for granted. On the contrary. A typical way to classify 4E views of cognition is to consider their respective attitudes towards the issue of representation and content. This issue concerns the extent to which cognitive processes involve the manufacture and use of naturally contentful intracranial states. Some of 4E theories are emphatically *non-representationalist*: they deny the existence—and the explanatory necessity of positing the existence—of mental representations understood as naturally contentful intracranial states. These theories do not only claim that the *traditional* notion of internal representation is superfluous; they claim that the explanatory appeal to *any* kind of natural content-involving intracranial structure is unnecessary and even ontologically unjustified. However, they do not deny the existence and explanatory relevance of linguistic or propositional content which is derived from linguistic practices. Other theories will rather endorse occasional non-representationalism: some basic or minor cognitive tasks do not require representationalist explanations, but other, more complex and evolved tasks surely require them (cf. for instance Andy Clark’s (1997) *representation-hungry tasks*). But these representations do not need to be static, complete, action-neutral or symbolic: they can be action-oriented, context-dependent, non-linguaformal, partial, modal, and so on.

Amongst these burgeoning theories, I will pay here special attention to *radical enactivism* and its rejection of representationalism. Since its inception, radical enactivism (or REC, for radical enactive cognition) sees itself as a revolutionary project in cognitive science, because of the

¹ It is also possible to understand “enactive” in a broader sense, encompassing the necessarily living and embodied character of cognition, as the works of Francisco Varela, Shaun Gallagher or Ezequiel Di Paolo show.

radical non-representationalist theorization it would propose of the 4E dimensions of cognition. More exactly, REC would be the equivalent, in cognitive science, of the revolutionary changes Einstein achieved for matter or Galileo for movement (Hutto and Myin 2017: preface and 46–47). Not less. Nevertheless, the problem with revolutionary claims in science is twofold. Firstly, they are neither necessary nor sufficient for bringing a real revolution. Claiming that you do a revolution does not entail by itself the occurrence of a revolution. Recent works have indeed begun to question the real revolutionary character of enactivism (Wheeler 2017). More broadly, the history of science is full of examples where thinkers who considered themselves as *orthodox* achieved what the posterity understood as scientific revolutions (take for example Copernicus), and full of examples of self-proclaimed revolutionary theories or sciences that have been finally classified as pseudo-sciences or scientific failures (see mesmerism or phrenology). Secondly, revolutionary claims like Hutto and Myin’s REC give the impression that the situation in cognitive science is clear: there are only two opposing alternatives (representationalism or content-involving cognitive science (CIC), as REC calls it, and REC), as if the opposition was perfectly real, and as if there were no other contenders in the field beyond mere variations around or between CIC and REC.

In what follows, I will notably show that REC and representationalism (or content-involving cognitive science (CIC), as REC calls it) are sharing important assumptions. It would be misleading to believe that since REC and CIC share these assumptions, these assumptions must be central and necessary for any kind of cognitive science. On the contrary: I will present positions that reject these assumptions. These positions have a *heuristic* value: assessing and defending their *scientific* value will not be my aim here. The assumptions which will be under question concern the relations between mental representation and mental content, and the status of intentionality in cognitive science. Contentless representationalism is a rejection of the identity between representation and content. Non-intentionalism is a rejection of the reality of intentionality as a basic, natural and intrinsic property of object-directedness. I will argue that contentless representationalism and non-intentionalism are raising challenges for REC and its revolutionary character at different levels: (1) contentless representationalism and non-intentionalism are alternative positions in cognitive science that are not mere variations around REC or CIC; (2) REC and CIC are rejecting contentless representationalism and non-intentionalism by sharing common assumptions about intentionality and about the equivalence between representation and content; (3) REC’s rejection of contentless representationalism is a threat to the intelligibility of its very own “nonrepresentational understanding of intentionality” (Hutto and Myin 2017: 15).

Section 1 presents contentless representationalism and the way REC rejects it by sharing important assumptions with CIC. Section 2 is

about REC's intentionalism and its ambiguities. I will notably discuss a reply made by Hutto and Myin (2017) to some objections expressed by Jean-Michel Roy (2015).

1. *Contentless representationalism and deflationism on mental representations*

Let us define representationalism (or CIC) as a commitment to the existence and theoretical relevance of mental representations, understood as material and intracranial entities, naturally endowed with content. But what is content? Classically, content is defined by truth conditions or satisfaction conditions. It may also be identified with abstract semantic entities like meanings, Fregean senses, possible worlds, modes of presentations, intensions, or propositions. These definitions are not identical, of course. But they all share one common methodological assumption: *the content of mental representations is defined by properties which also figure in the definition of the content of public or linguistic representations, namely meaning, reference, truth-conditions, or satisfaction conditions.* The only difference being that unlike the content of public/linguistic representations, the content of mental representations is natural, and thus fixed by natural properties and relations. Daniel Dennett clearly expressed the basic prejudice at the core of this standard view of the content of mental representations:

Whatever *mental* representations are, they must be understood by analogy to *nonmental* representations, such as words, sentences, maps, graphs, pictures, charts, statues, telegrams, etc. (Dennett 1978: 189, author's emphasis)

As stipulated theoretical entities, mental representations are nevertheless modeled on a kind of ordinary public representation: linguistic representations.

This approach to mental content has been discussed for some time now. Examining the way this standard view has been questioned by various philosophers will lead us to a basic problem that also concerns the eliminativism about content and mental representation developed by REC.

From his classical 1989 book *Meaning and Mental Representation* to more recent papers,² Robert Cummins has argued that there is no need to suppose that “representation” and “content”, in cognitive science, should have the same explanatory role and the same properties that these terms have in folk psychology. “Content”, for instance, should not be confused with meaning (the latter one being a property of natural languages), defined in terms of truth conditions or satisfaction conditions, or associated with reference. Separated from these linguistic phenomena, the problem of naturalizing content is not the problem of

² See especially the paper co-authored with Martin Roth “Meaning and content in cognitive science”, reprinted in Cummins (2010: chap. 11).

finding and defining a natural semantic *relation* between intracranial states and environmental states of affairs. Content might just be identified with functional or inferential role. It is therefore a *petitio principii* to assume that the content of mental representations—whatever it is—*must* have properties similar to the ones common-sense associates with content, like for instance meaning or truth conditions. As Stephen Stich (1983) noted a long time ago, content ascriptions are irreducibly vague, context-sensitive, and observer relative. For these reasons, content *as it is ascribed in folk psychology* may have no place in scientific psychology. But that does not mean that the concept of “content” has no role to play in cognitive science (Stich 2009: 204), since it is invoked in successful theories (Stich 1996: 199). The place of “content” in cognitive science can be secured not by naturalizing it, but by defining what a “successful” theory is, not only from internal features, but also external ones (sociological success, entrenched habits of problem-solving,...). According to Cummins and Stich, the content that is invoked in cognitive science must be kept apart from meaning, truth-conditionality, and more broadly linguistic content. If we follow this strategy of emancipating content in cognitive science from folk or linguistic accounts of content, it is a mistake to argue for the elimination of mental representations from the fact that the content associated to representational entities by the common-sense conception of “mental representation” has no scientific reality or cannot be naturalized.

We find here a classical problem for any eliminativist argument that starts from a classical common-sense characterization of the property to be eliminated. As Steven Stich and Stephen Laurence say:

Those arguments typically begin by describing some feature or cluster of features that are important or essential for intentional states, *on the commonsense account of these states*. The arguments then try to show that respectable scientific theories cannot accommodate states with the features in question. (Stich and Laurence, in Stich 1996: 178, emphasis mine)

Eliminativist strategies generally start from a descriptivist model of reference, according to which the reference of a term is determined by the cluster of descriptions associated with this term, for instance descriptions coming from commonsense. The main assumption of a descriptive theory of reference is that if a theoretical concept *C* refers, it refers to whatever is picked out by the description associated with it in the theory (Devitt 2009: 46). If nothing satisfies that description, there are good reasons to think that the concept does not refer to anything. However, things are different if one endorses a causal theory of reference: a concept might refer to something whose properties are not the ones mentioned by the intension of the concept, especially if this is a scientific concept whose intension is initially defined from commonsense. The meaning of scientific terms may not be fixed by their intension; it may be determined by the nature of their referents, as Kripke and Putnam taught us a long time ago. A challenge appears for any eliminativist position, including REC’s eliminativism on content and

mental representation: what are the conditions in virtue of which one can say that some entity or property (here: mental representation) does not exist, rather than say that it exists, although it is very different from what one thought and thinks about it? What are the conditions in virtue of which some term does not refer to anything, rather than referring to something that is very different from what the descriptions associated with it prescribe?

This problem should not only be faced by eliminativists on mental content who assume that mental content is *necessarily* truth-conditional, propositional or conceptual. REC's acceptance of the fact mental content does not necessarily require truth-conditionality or intentionality must not hide its commitment to a more basic assumption whose refusal is sufficient for blocking REC's eliminativist strategy: the assumption that mental representations necessarily have contents (whatever the properties associated with content) and referential properties. Refusing to endorse *one* specific classical conception of content (the conception according to which mental content is necessarily propositional or truth-conditional) does not mean one is not under the grip of a more basic conception of content whose parochial and pre-theoretical character may also be questioned.

Let us call [CONTENT-TRUTH] the claim that natural mental content is necessarily truth conditional. This claim may be endorsed by both proponents of natural mental content and eliminativists about natural mental content. As said before, Cummins and Stich have convincingly argued that [CONTENT-TRUTH] is not the only account of mental content which is available in cognitive science. This non-necessity is already sufficient for criticizing eliminativist strategies on mental content which assume that mental content does not exist because there are no truth conditions in intracranial mental structures or in perceptual experience. Failure to naturalize (so that elimination follows) truth-conditional or meaningful content does not imply failure to naturalize (and thus the elimination of) another kind of content. REC also refuses that [CONTENT-TRUTH] is the only existing account of content in cognitive science. More modest accounts of content have been proposed by philosophers such as Tyler Burge, who identifies content with accuracy or veridicality conditions (Burge 2010). REC is also targeted at these accounts. Since their first book, Hutto and Myin argue that neither informational theories nor teleosemantics are able to provide a satisfactory non-intentional explanation of the emergence of semantic properties (be they a matter of truth conditions, satisfaction conditions, accuracy conditions; be they conceptual or non-conceptual; be they propositional or non-propositional): either these theories beg the question by already coming with intentional notions, or they merely deliver covariation and indication, which are not sufficient for giving semantic or representational content. These failures to naturalize content (what REC also calls "The hard problem of content") entail that representationalism has no foundations in the naturalistic ontology proponents of

representationalism generally assume. Unable to be integrated in the naturalistic ontology it claims to be a part of, the representationalist program would be “plagued with toxic debt, financed by loans it cannot pay back” (Hutto and Myin 2013: 160). Since mental content has no place in a naturalistic ontology, there are good reasons to think it does not exist as an entity conveyed or produced by natural processes, including subpersonal and intracranial ones. In addition, non-representational means and models are already available and plausible for explaining basic cognitive phenomena. And, according to REC, when it comes to contentful cognition (thoughts, imaginings or reasonings), contents are not natural or subpersonal contents: they derive from the integration of cognitive agents in socio-cultural practices.

Hutto and Myin claim that respectable naturalistic theories cannot accommodate naturally contentful cognitive states (mental representations), so that these states should be theoretically eliminated. REC assumes here—as a *petitio principii*—what I will call here [REPRESENTATION-CONTENT-REFERENCE]: the claim that mental representations *necessarily* have content and entertain semantic relations with their referents.

In order to clearly define the pivotal role of this (resistible) assumption by REC, let us consider here a reconstruction of Hutto and Myin’s strategy in their 2013 book; this strategy has not changed in their 2017 book:

- P1. In the representationalist ontology, the subpersonal and intracranial phenomena named “mental representations” naturally (or intrinsically) have contents (truth-conditions, satisfaction-conditions, or accuracy conditions) and (by implication) reference³;
- P2. There is no natural (or intrinsic) content at the level of subpersonal and intracranial phenomena;
- C. Subpersonal and intracranial mental representations, as they are conceived by the representationalist ontology, do not exist.

P2 is defended from an examination of the failures of projects of naturalizing mental content. A reply from representationalists may consist in a criticism of P2. Representationalists can argue that content has been naturalized (see Milkowski 2015), or argue that the fact it has not been naturalized yet does not entail it is not naturalizable (and thus existing).

Even if P2 is correct (“there is no natural content at the level of intracranial and subpersonal phenomena”), the proponents of representationalism may also object to the general argument by refusing P1 as it is stated by REC. P1 is the description of an alleged consen-

³ I here leave aside the theoretical possibility of representationalist theories that would hold that mental representations do not have content, but nevertheless have reference. For the all representationalist and non-representationalist theories I consider here, content determines reference. Content and reference stand or fall together.

sus: in the representationalist community, everyone would assume that mental representations are endowed with natural content. This corresponds to [REPRESENTATION-CONTENT-REFERENCE]. Even if REC questions the existence of mental content and, therefore, of mental representations, it does not disagree with this characterization of mental representations: they (are supposed to) have content. “Content” is part of the intension putatively fixing the reference of the concept “mental representation”. Representationalists can argue that mental representations never have natural or intrinsic content (in this case they would agree with REC: mental content derives from social and cultural practices), but they can also argue that mental representations do not have content *at all*, so that P2 has no consequences at all on the existence of mental representations. What P1 describes is a situation in which commonsense provides the intension from which mental representations in cognitive science should be conceived: they have content and referential properties, just like sentences, pictures or diagrams. But if one denies that mental representations have content, P2 has no consequence at all for the existence of mental representations (or for the reference of the term “mental representation”).

But are there representationalists that *seriously* deny P1 and thus reject [REPRESENTATION-CONTENT-REFERENCE]? Of course there are. Outside of the context of cognitive science, eminent philosophers such as Nelson Goodman (1968) or Dan Lloyd (2003) have long questioned the idea that the property of representation is a relational property, and not (for instance) a monadic property. Let us consider closely the case of Noam Chomsky.⁴ For a long time now, Chomsky is convinced that folk and philosophical notions like “content”, “intentionality” and “reference” have no place at all in the naturalistic framework dedicated to the understanding of cognitive faculties (Chomsky 2000: 21–23). This is of course much more radical than asserting, like Cummins, that “content”, in cognitive science, has a different sense than “content” in folk psychology and semantics. As Chomsky writes:

The central problem that troubles me is this. I do not know of any notion of ‘representational content’ that is clear enough to be invoked in accounts of how internal computational systems enter into the life of the organism. And to the extent that I can grasp what is intended, it seems to be very questionable that it points to a profitable path to pursue. (Chomsky 2003: 274)

Nevertheless, for Chomsky, the concept “mental representation” should be retained in cognitive science, but “content”, “reference” or “intentionality” have to be purged from its intension. That is, Chomsky emphatically refuses [REPRESENTATION-CONTENT-REFERENCE]. Mental representations are individuated from their role in computational processing. The functional roles of mental representations are here related to properties that have nothing to do with content, ref-

⁴ In cognitive science, Ray Jackendoff (1987) would have been another possible example.

erence, or intentionality. Their important properties are formal or syntactic. These representations do not mean or represent anything; defining their reference is of no scientific interest. One of the reasons of Chomsky's eliminativism on reference, content and intentionality is related to the difficulties of *individuating* the objects of reference, be they actual or purported (Chomsky 2003: 273). The use of intentional expressions such as "refers to" or "means" can be preserved for informal expressions of a computational theory, but intentional expressions play no role in the computational theory itself. The computational theory makes use of the concept "representation", but this concept is not about any relational entity. This does not exclude the introduction of a technical notion of reference in order to explain the syntax of mental representations (Chomsky 2000: 202 n.6):

The internalist study of language also speaks of "representations" of various kinds, including phonetic and semantic representations at the "interface" with other systems. But here too we need not ponder what is represented, seeking some objective construction from sounds or things. (...) Accessed by performance systems, the internal representations of language enter into interpretation, thought, and action, *but there is no reason to seek any other relation to the world, as might be suggested by a well-known philosophical tradition and inappropriate analogies from informal usage.* (Chomsky 1995: 53, my emphasis).

"Informal usage", here, means the very widespread tendency to embrace a linguistic model of mental representations, assuming they have semantic content or reference, like daily linguistic products. Chomsky's "contentless representations", as we may call them, entertain functional relations with external phenomena: they occur when and only when the organism interacts or deals with these external phenomena. In this sense, a "number-representation" is a representation of a different functional type than a "face-representation", but is not to be defined as a representation *of* an external item (Chomsky 1995: 52).

But what would be the utility and the plausibility of this notion of "contentless representation" for cognitive science? A double answer might be proposed to this question, one answer justifying why the notion of "contentless representations" might deserve to be used for naming some specific processes, the other one justifying how it is fruitful to see those representational processes as being *contentless*. Firstly, using the notion of "representation" for labeling a subpersonal process is here a way to underline the fact this process has specific cognitive properties. This subpersonal process functionally contributes to the realization of the cognitive abilities of the system (or person) under study, but not only. Indeed, these abilities are those which are exercised when the system deals with environmental objects, properties or states of affairs in tasks such as perceiving, memorizing, or understanding. Representation-talk is a way for the theorist to individuate and classify these cognitive processes, underlining their functional relations with external objects. Labeling a structure as a "face-representation", "phoneme-

representation”, or “space-representation” is underlining the fact this structure plays a role in the relations the cognitive system entertains with faces, phonemes or space. Nevertheless, since individuation is not definition, that does not entail these processes entertain intrinsic semantic relations or reference with these objects, or that they specify them in virtue of some content. The relations between these structures called “representations” and external objects are not semantic or contentful, but they are more than mere causal relations, because these structures are theoretically individuated as playing a key functional role in the cognitive processing in relation with these objects and properties. Still, the definition of the formal and syntactical properties of those representations is sufficient for studying their causal role in cognitive processing. Secondly, by considering these representations as being “contentless”, one does not introduce into the theory notions (as “content”, “reference” or “truth conditions”) first proper to philosophy of language, and which immediately raise the thorny issues of naturalization and causal efficiency. How could natural facts generate semantic or contentful properties? How could semantic or contentful properties be semantically efficient? Those vexed issues disappear for those who dispense with content. Content-ascription can play some auxiliary role in the informal presentation of the computational theory, but not within the computational model itself.

As said in the introduction, my aim here is not to assess or to defend the scientific plausibility of this marginal position in cognitive science (see Rey 2003; Jacob 2010; and Egan 2003 for some existing assessments).⁵ I prefer to focus on the arguments in virtue of which this position is (unsurprisingly) rejected by both REC and its classical opponent.

Facing the suggestion of divorcing mental representations from content, intentionality and reference, *both* classical representationalists and REC might reply that “contentless representation” is an oxymoron. And, actually, this is REC’s answer (see Hutto and Myin 2013: 84): if something does not have content or reference, why persisting in calling

⁵ More broadly, I do not assume here that there is or must be *one* right (technical, scientifically respectable) way to conceive content and mental representation in cognitive science. This realistic assumption might be shared by many representationalists (who believe in the reality of mental representations) and anti-representationalists (who consider that there are *facts of the matter* in virtue of which there are no mental representations); but it is optional. My instrumentalist and more precisely pragmatic conception of inquiry in cognitive science leads me to think that it is only from local, well-identified cases of cognitive explanations that we can raise questions about the (local) usefulness or acceptability of “representation”, “content” or “non-representation” talk. This important issue is quite orthogonal to the issue of knowing whether REC and representationalism share common elements, or whether REC is a coherent position in itself. Still, it informs the pluralistic perspective I take in this paper when I argue that we should not believe it is *necessary* to use folk concepts of “content” and “representation” in cognitive science. I thank one anonymous referee for having invited me to clarify this point.

it “representation”? This reply, at least made by REC, assumes that representation and content stand and fall together. This assumption may be disputed: why should cognitive scientists absolutely respect common sense conceptions of “representation”? Take terms such as *matter*, *mass*, *life* or *length*, as they are defined and used in contemporary science. That is, consider them as *scientific concepts*. Matter, in contemporary physics, is not necessarily impenetrable and solid. Mass is interchangeable with energy. Life does not necessarily require replication or reproduction. Length is a function of relative velocity. Are these definitions invalid or unacceptable on the ground that they do not match with ordinary or commonsense understanding of these terms? Is there any obligation for scientists to forge new terms because of the confusions these new uses of entrenched terms can generate? Is there any *naturalist* philosopher that wants to subject theories in chemistry, biology or physics to the authority of ordinary language? Is there any naturalist philosopher that would argue that psychology and cognitive science—unlike natural sciences—must be subjected to this authority, and that will thus reject the possibility there might be contentless mental representations? This is a first set of questions that may follow from REC’s dismissal of contentless representationalism.

Pursuing my project of focusing on REC’s treatment of contentless representationalism, let me note another point (and problem). Hutto and Myin sometimes reduce Chomsky’s position to the following claim: since “mental representation” does not pick out any contentful or referential state, it picks out *any* property of brains that we have not exactly defined yet. “Mental representation” would be a mere label that picks out whatever, in the end, best characterizes what does the work in cognition. The obvious reply REC proposes to this definition is that, in the end, mental representations would be everywhere, the concept being applied to any kind of mediating state (Hutto and Myin 2019: 8).

Hutto and Myin’s objection is instructive: they are warning non-orthodox representationalists who want to give up content (as Chomsky) that their representationalist ontology is under the threat of pan-representationalism. Mental representations would be everywhere, in any mediating (and contentless) physical state. Hutto and Myin’s objection assumes that *content is the only good criterion for preserving an intelligible form of representationalism*. But, of course, because of content, this form of representationalism would then face the “Hard Problem of Content” and hence the prospects of elimination. Put otherwise, Hutto and Myin invite all representationalists to endorse a specific view of mental representation, so that eliminativism about content can automatically generate eliminativism on mental representation. The non-orthodox, deflationary contentless representationalist would be trapped in a dilemma, between the hard problem of content and pan-representationalism:

The deflationist’s dilemma is this: either retain mental contents and their

troublesome properties or let go of mental contents and offer a theory of mental representation that is indistinguishable from non-representationalist accounts. (Hutto and Myin 2019: 22)

The upshot of this strategy is twofold: on the one hand, in order to criticize representationalism, REC is forced to endorse a basic tenet of classical representationalism (mental representations have content and reference), and might more broadly share with classical representationalism an occasional reference to common sense as a landmark in the characterization of the acceptable and unacceptable posits of cognitive science (indeed, as seen above, according to REC there is no representation without content, and “contentless representation” is an oxymoron). On the other hand, by addressing a dilemma to deflationists, one can wonder whether Hutto and Myin are not cutting off the branch they are sitting on regarding their very own conception of *intentionality*. Indeed, if there is no natural content and therefore no mental representations, why should there be intentionality? If intentionality is contentless, how should and could one distinguish intentionality from non-intentional phenomena? Is there not now a lurking problem of pan-intentionalism? This will be the topic of the next section.

2. *Non-representationalism and intentional realism: How low can you go?*

Any textbook in philosophy of mind will reveal that the property of “intentionality” is a fundamental property of minds. But what is intentionality? Apart from the basic idea of intentionality as a *relation* between minds and world, consensual definitions are hard to find; metaphors abound. Linguistic models are often convoked for defining and individuating intentional states: reference, representation meaning and content would be proper to intentional states. Here is for instance John Searle:

Intentional states represent objects and states of affairs in the same sense of ‘represent’ that speech acts represent objects and states of affairs. (Searle 1983: 4)

Associating intentionality with representation is the core of the *representational* theory of intentionality (Cummins 1989: chapter 1; Morgan and Piccinini 2018). According to it, the contents of intentional states are mental representations. In order for S to be intentionally related to O, there must be a mental representation of O in S. This works for intentional relations with objects, but also with propositions (Field 1979; Fodor 1985). Representationalist theories of intentionality often equate intentionality with representation and content. Intentional states would be contentful states, or states that represent objects. The problem of intentionality would be the problem of explaining how some entities can represent or stand for other entities (Stalnaker 1984: 6). This ambiguity or even confusion between representation and inten-

tionality has some theoretical advantages: it easily gives the impression that intentionality can have a causal role in the physical world, in virtue of the physical and formal properties of its representational vehicles.

Because of its anti-representationalism, REC cannot be suspected to endorse a form of representational intentionalism. But there is a more basic tension than the tension between representationalism and non-representationalism when one wants to offer an account of intentionality. It is the tension between intentional realism (or intentionalism) and non-intentionalism. And, from the perspective of this basic tension, REC and representationalism are in the same boat: they endorse (different versions of) intentionalism—like other enactive theories of cognition (Varela 1992, Thompson 2007, Gallagher 2017, Noë 2004) I will not discuss here.

For REC, basic cognition is a matter of embodied engagements responding to wordly offerings or informations in the environment (Hutto and Myin 2017: 130). These responses do not involve contents; but they must be explained, and in particular their *connecting* properties with the world:

Anyone who claims that cognition is entirely a matter of contentless computations—for example anyone who allows that content falls out of the equation entirely, and offers no successor notion—will be unable to explain how organisms relate to and connect with targeted aspects of their wordly environments. Any theory of this extremely austere sort will be woefully ill-equipped to explain the array of findings that give us reason to think that cognitive activity is deeply influenced by E-factors. (Hutto and Myin 2017: 59)

Getting rid of content is not sufficient for being revolutionary. You also need to provide an alternative (that is, contentless) story to a classical question: how do organisms relate to aspects of the world? Answering to this question is pressing, especially if one holds that cognition is *embodied, extended and embedded*: how can agents relate to parts of the world that are influencing and even constituting their cognitive operations? Facing this challenge, the own *explanans* of REC is *intentionality* or more precisely *Ur-intentionality*, the “most primitive form of intentionality” (2017: 96). Intentionality is thus, for REC, the basic operator that will ground a 4E approach to cognition:

REC questions whether, on close inspection, there is a need to posit any kind of content at the basement level of cognition in order for the sciences of the mind to do their fundamental explanatory work. *On the positive side, REC recommends getting by with something less—an alternative, contentless notion of intentionality.* In short REC avoids a host of intractable problems—most prominently the HPC [*Hard Problem of Content*] and the problem of mental causation—by *sticking with the idea that organisms target chunks of the world* without assuming semantic contents make any causal or other explanatory contribution when it comes to saying how such targeting is possible. (Hutto and Myin 2017: 60, my emphasis)

To put it in a nutshell: “basic minds target, but do not contentfully represent, specific objects and states of affairs” (Hutto and Myin 2017: 130). Already in their 2013 book, Hutto and Myin stated that organisms exhibit “intentional directedness” towards aspects of their environment (2013: 81) (see also Hutto 2008: chapter 3 on the differences between intentional attitudes and propositional attitudes). This intentionality is a property of organisms, and not of mental or physical states inside of these organisms.

REC therefore endorses a form of intentional realism (or intentionalism): intentionality is a real, natural and intrinsic property of organisms. It is not a matter of observation, description, and interpretation. Organisms display Ur-intentionality independently of what one may think or say about them, and independently of their possible inclusion in socio-cultural practices. The “reality” of intentionality amounts to its being natural, and naturalizable from the resources of teleosemiotics. Intentionality is neither magic nor given: intentionality has been shaped through ontogenetic and phylogenetic history (Hutto and Myin 2017: 108, 130; Hutto and Myin 2013: 111).

This definition of intentional realism is very distinct from classical definitions and forms of intentional realism, which systematically associate intentionality with content and representations. Being realist about intentionality would be being realist about the existence of mental content, propositional attitudes and mental representations (see for example Jacob 1997: chapter 1). Conversely, anti-realism about intentionality would be the idea that there is no “fact of the matter about what a person (or a person’s mental state) *really means*” (Dennett 1987: 294). Since REC urges us to divorce intentionality from meaning and representation, its form of intentional realism must not involve content, intentionality, meaning or representation in the very definition of intentionality. What is *real*, in REC’s intentionalism, is the property of “aiming at” or “pointing towards” worldly offerings (Hutto and Satne 2015: 530, note 7). Indeed, if linguistic landmarks and representationalist descriptions are refused for defining intentionality, other metaphors can be used. Metaphors such as “aiming at”, “targeting” or “pointing” are supposed to suggest the core of intentionality, echoing the etymology of the word (cf. the latin verb “intendere”, “aiming at something”). Even though REC gives up representationalist and semantic readings of aboutness, it keeps the idea of intentionality as a real power of organisms. Being realist about intentionality is considering that the aboutness of mental states is the result of their having an (*intrinsic*) *power of aiming at, or being directed to, objects*. There is a pointing-beyond-itself which is characteristic of some entities.

For non-realism about intentionality, there is no intrinsic intentionality, be it contentful or not. Nature does not produce meaning or content (on this point, Hutto and Myin are right), but not only: the aboutness of mental states and organisms is not a matter of their being *intrinsically* aimed or targeted at something (it is here that Hutto and

Myin disagree with non-realism about intentionality). More precisely, one version of non-realism will argue that directedness to objects is so pervasive and trivial in nature that it is insufficiently interesting and specific for providing that philosophical property intentionalists like Hutto and Myin name “intentionality” (Rosenberg 2013: 2015). Intentionality can for instance rather be a property of the linguistic *articulation* of cognitive and mental attitudes. Another version will argue that object-directedness is always observer-dependent, and thus never intrinsic (some versions of Maturana-inspired autopoietic enactivism endorse this idea).⁶ One proof of the proximity between REC and its classical representationalist “opponent” is their common reaction to non-realism about intentionality: incredulity. As an example of the important foundational role it gives to intentionality as a real property, one may consider REC’s criticism of *neo-pragmatist* accounts of intentionality, which hold that all intentional properties are derived from linguistic and social practices. For REC, the denial of natural intentionality entails that neo-pragmatism has no means “of accounting for the kinds of intelligent thinking that are needed for explaining participation in the relevant socio-cultural practices” (Hutto and Satne 2015). In short, a natural and real form of intentionality would exist, and would be the *only* way of explaining basic forms of intelligent thinking—as if other (non-intentional) answers and proposals were *a priori* unavailable or considered as dead ends or non-starters.

The intentional realism (or intentionalism) of REC can raise suspicion from two very different sides: for representationalist intentionalists, REC’s intentionality is not intentionality, because it does not involve representational properties or content. For non-realism about intentionality, REC’s intentionality is so low and basic that it is virtually indistinguishable from non-intentional relations (we here meet the dilemma presented in the first part of the paper). In order to develop this last point, two supplementary remarks must be made.

First, REC’s intentional realism embraces the classical picture of intentionality as a relation to *objects*. Intentionality is *object-directedness*. If the frog is intentionally related to the world, it is because it is related to objects, and not mere things: it is related to entities that have a specific behavioural profile (Hutto and Myin 2017: 111–112). It aims at objects, because *object-directedness* is defined, since Brentano, as a core property of intentionality.

Second, REC is at pains to insist on the fact that being intentionally directed at something is not only being disposed to do something; it has a normative dimension. More exactly, what is intentionally targeted is normatively fixed by past interactions between organisms and their environment (Hutto and Myin 2017: 112). The natural attunement between organisms and their environments in the past not only structures the profile of an organism’s current tendencies for response, it

⁶ See Abramova and Villalobos (2015).

normatively fixes what is intentionally targeted, in complicated ways across multiple spatial and temporal scales.

These two remarks can help us to understand how Hutto and Myin were able to respond to an important objection regarding the reality of their own “Ur-intentionality”. As Roy (2015) was one of the first commentators to notice, the main challenge of REC’s intentionalism is to show how a non-semantic and contentless relation between organisms and parts of the world should nevertheless be seen as an intentional relation. One can address this challenge to REC without being a (covert) representationalist.⁷ Indeed, one can reject the existence of Ur-intentionality not because it is contentless or devoid of semantic properties, but because it seems to correspond to a relation which is so basic and pervasive that it is useless and misleading to call it “intentionality”. There are, to be true, “biologically forged mind-world connections”, and indeed, they do not have “to be characterized in semantic terms” (Hutto and Myin 2017: 108). But, more fundamentally, why turning these “natural involvement relations” (Godfrey-Smith 2006: 60) or “natural attunement relations” (Roy 2015) into intentional relations? If they are so basic, why seeing them as intentional? (and not: “if they are not semantic, why seeing them as intentional?”). There are two answers provided by Hutto and Myin: a broad answer, already mentioned above, arguing that it is *necessary* to see these relations as intentional relations if one wants to understand how cognitive organisms connect to their environment. In line with intentionalists, REC considers that the relation between organisms and their world is a scientific question, and that intentionality is the (scientific) answer to it⁸. And there is a narrow, and more technical answer, straightforwardly directed to Roy’s question (Roy 2015: 123), who asked why the directedness organisms have towards elements of their environment should be understood in terms of intentional directedness to objects, and not merely in terms of behavioural attunement. To this objection, Hutto and Myin replied in their 2017 book:

It is because REC casts Ur-intentionality in normative terms that it does not equate basic intentional directedness “to a sort of property of natural attunement and thus loses its connection with... objectivation” (Roy 2015). (Hutto and Myin 2017: 112)

It is thus the normative character of the targeting that grounds the objective dimensions of intentionality and more broadly that preserves the place of intentionality in a naturalistic framework. Indeed, as Hutto and Myin say, “it is precisely because REC makes room for at

⁷ Contrary to what Hutto and Satne (2015: ft.7) suggest, when they say that the ones who reject Ur-intentionality proceed by defining intentionality in terms of content. That is an incorrect characterization of all their possible opponents.

⁸ This can be contrasted with Dennett’s non-intentionalism: “The phenomena of intentionality are both utterly familiar—as salient in our daily lives as our food, furniture, and clothes—and systematically elusive from scientific perspectives” (Dennett 2013: 64).

least this much normativity that it differs from the eliminativist, strict naturalist approaches (...)" (2017: 112). Normativity makes for objectivation, which is the defining feature of intentionality.

I would like now to discuss Hutto and Myin's reply to Roy: is this reply sufficiently clear and coherent for dispelling non-intentionalist worries?

(1) *Normativity and objectivity.*

Let us accept that natural normativity is real, and not observer-dependent. Let us accept, by charity, that "there are historical facts about what ancestral organisms interacted with in their environments that shaped, and currently constrain, the response profiles of members of any given species" (Hutto and Satne 2016). How is that sufficient for positing intentionality as the power of targeting objects? That is, how should the presence of normativity be sufficient for providing object-directedness, which is seen by REC as the defining feature of intentionality? Representationalist intentionalism will ground object-directedness on representational content (content is the aspect under which a thing is apprehended, and is thus an object for a mind); this road is not available to REC, but can normativity do the main work here?

There are many devices in nature (hearts, kidneys, intestines...) that function normatively, if we stick to REC's characterization of normativity, but that does not make them devices which are targeted towards *objects*. Organs selectively respond to viruses and bacteria; does this mean they are targeted to viruses or bacteria as objects, and are thus intentional systems? A threat of bloat (or pan-intentionalism) appears: we met it at the end of section I concerning contentless representation (if a mental representation is just a mediating structure, why should not the world be full of mental representations?). We meet it now for intentionality: if intentionality is a capacity of living systems to target (with possible error) objects, this capacity being subject to history and norms, then why are not sunflowers, kidneys, bacteria and intestines bearers of intentionality?

If REC challenges Chomsky's contentless representationalism by arguing that it entails that mental representations are everywhere, this challenge should also concern their own position concerning intentionality. Why being so demanding for content and representation (nothing in nature can constitute them) and so liberal for intentionality (it is enough, for intentionality, to be made out of normative relations)? The objection according to which *objective transitivity* (being directed to objects) is too broad and general for defining intentionality as a defining feature of cognitive or mental systems is not new: it was already addressed by Edward Titchener (1909: chap. 2) to Brentano's idea that intentionality would be the defining mark of mental phenomena.

REC is not alone facing this problem of pan-intentionalism. And it might transform this problem into a basic commitment: there is inten-

tionality as soon as there is life. REC can here join forces with some versions of *autopoietic* enactivism which do not hesitate to see basic forms of intentionality in very basic living systems, like bacteria (Varela 1992).⁹ For REC too, simplest life forms are capable of an intentionally directed responding and directedness (Hutto and Myin 2013: 36), although REC is suspicious towards the use of concepts like “meaning”, “contents”, “sense-making” and “significance” for describing the responses of simple living systems (2013: 35). “Basic interest-driven ways of responding” would be the right platform for understanding how mentality can be intentionally directed (2013: 36). Agreeing with these intentionalist versions of autopoietic enactivism, and accepting that intentionality is present as soon as there is life, REC would also not be very far from Tecumseh Fitch’s model of *nano-intentionality* (Fitch 2008): there is already an intrinsic goal-directedness inherent in the behaviour of living eukaryotic cells, and more precisely a goal-directed capacity to respond adaptively to novel circumstances, by arranging and rearranging molecules in a locally-functional manner, based on past history. In doing so, cells autonomously arrange their form in such a way as to optimize their ability to perform certain quite specific functions. Of course, goal-directedness is not object-directedness. It remains to be seen whether Fitch’s nano-intentionality exhibits the essential feature of objectivation which is classically associated with intentionality. Be that as it may, by going biological about intentionality, REC would meet again the threat of pan-intentionalism: there would be intentionality everywhere there are basic biological systems such as cells and bacteria. Intentionality would become a universal generic property of living systems, and we would need many specific distinctions inside of the realm of living beings in order to account for the differences and the relations between, for instance, the intentionality of bacteria, the intentionality of protozoa, the intentionality of an organ, and the intentionality of an organism. What would be the relations between, say, the Ur-intentionality of the organism and the intentionality of the billions of cells it is made of, and the relations between this organismic Ur-intentionality and the intentionality of propositionally contentful mental states? Moreover, if REC considers that intentionality (as object-directedness) is already present in cells, it owes us a new definition of the role of intentionality in cognitive science. How much would intentionality be a specific cognitive property or phenomenon studied by cognitive science (and not biology)? Does it also entail that artificial creatures cannot exhibit intentionality or harbour states having intentionality?

⁹ See Abramova and Villalobos (2015) for a non-intentionalist version of autopoietic theory: even contentless, intentionality as directedness is a case of anthropomorphic projection which is ascribed by the observer to the organism.

(2) *The place of intentionality in nature.*

By getting rid of content, REC escapes the problem of explaining how semantic properties can be causally efficacious, but it does not escape the problem of accounting for the causal role of intentionality in nature. If intentionality is a real-intrinsic property of organisms that deserves to figure in our best theories of cognition, the least we may ask is this: what difference does this property make? REC will answer by saying that intentionality is what allows organisms to be related to specific aspects of their environment. Very well, but an explanation of how intentionality *works* for achieving that role is still required. The causal role of intentionality cannot be, of course, a matter of mental representations; REC will rather appeal to the instantiation of intentionality in behavioural dispositions shaped by evolutionary history (Hutto and Myin 2017: chap. 5). However, why should intentionality be something more than a *name* for qualifying these dispositions? It is not because you may describe organisms having dispositions with the use of the intentional idiom (“aims”, “targets”, “is directed towards”...) that organisms *become* intentional agents for real. Organisms have propensities and dispositions, products of a long history and learning. Intentionality might not be distinct from them; it might just be a way to describe these propensities and dispositions, and especially their objects-involving manifestations.

From these two objections, we can now paraphrase the dilemma REC stated for contentless representationalism as follows, and target it towards the challenge of explaining the nature and the causal role of Ur-intentionality: “either retain mental contents and their troublesome properties or let go of mental contents and offer a theory of intentionality that is indistinguishable from non-intentionalist accounts”. In sum, Hutto and Myin’s radical view of intentionality, stripped of all its representationalist apparatus, allows us to inquire over the very idea of intentionality: what is this special capacity or power of minds (or organisms) to target some objects? Is there any? Should there be any?

Conclusion

REC’s criticism of representationalism assumes that mental representations (if any) have content (whatever its properties) and reference. This assumption is refused by some varieties of representationalism, as contentless representationalism, which is therefore left untouched by the anti-representationalist strategy of REC. Moreover, this assumption is shared by REC’s best enemy: content-involving cognitive science. In order to defuse the plausibility of contentless representationalism, REC has expressed a dilemma; but this dilemma can also be applied to its own positive model of (non-representational) intentionality. Like CIC and more broadly classical cognitive science, REC considers that intentionality is a basic property of cognitive systems.

However, the way it characterizes intentionality seems insufficient for distinguishing intentionality from non-intentional relations, unless one accepts intentionality is present in every living system.

Up to now, there has not been a real debate between REC and classical representationalism, perhaps because proponents of the latter one believe proponents of REC live in a very different world. Hence disdain or incredulity. But they are not living in a different world: on some aspects, REC and classical representationalism are sharing important presuppositions that may be at the origins of some of the problems they respectively face. REC is not necessarily *revolutionary*: it retains some crucial assumptions shared with CIC, as the equation between representation and content, and intentionalism. Nevertheless, it may well be *radical*, in the sense that it forces us to focus on these foundational issues more clearly and demandingly than before.

References

- Abramova, K. and Villalobos, M. 2015. "The apparent (Ur-)intentionality of living beings and the game of content." *Philosophia* 43/3: 651–668.
- Chemero, A. 2009. *Radical Embodied Cognitive Science*. Cambridge: MIT Press.
- Chomsky, N. 1995. "Language and nature." *Mind* 104 (413): 1–61.
- Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, N. 2003. "Reply to Egan". In L. Antony and N. Hornstein (eds.). *Chomsky and His Critics*. Oxford: Blackwell: 268–274.
- Clark, A. 1997. *Being there. Putting brain, body and world together again*. Cambridge: MIT Press.
- Cummins, R. 1989. *Meaning and Mental Representation*. Cambridge: MIT Press.
- Cummins, R. 2010. *The World in the Head*. Oxford: Oxford University Press.
- Dennett, D. 1978. *Brainstorms*. Cambridge: MIT Press.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge: MIT Press.
- Dennett, D. 2013. *Intuition Pumps and other Tools for Thinking*. New York: Penguin.
- Devitt, M. 2009. "On determining what there isn't." In D. Murphy and M. Bishop (eds.). *Stich and His Critics*. Oxford-Malden: Blackwell: 46–61.
- Egan, F. 2003. "Naturalistic inquiry: where does mental representation fit in?" In L. Antony and N. Hornstein (eds.). *Chomsky and His Critics*. Oxford: Blackwell: 89–104.
- Field, H. 1978. "Mental representation." *Erkenntnis* 13 (1): 9–61.
- Fitch, T. 2008. "Nano-intentionality: a defense of intrinsic intentionality." *Biology and Philosophy* 23 (2): 157–177.
- Fodor J. 1985. "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum." *Mind* 94 (373): 76–100.
- Gallagher, S. 2017. *Enactivist Interventions: Rethinking the Mind*. Oxford: Oxford University Press.

- Godfrey-Smith, P. 2006. "Mental Representation, Naturalism, and Teleosemantics." In G. MacDonald and D. Papineau (eds.). *Teleosemantics*. Oxford: Oxford University Press: 42–68.
- Goodman, N. 1968. *Languages of Art*. Indianapolis: Hackett.
- Hutto, D. 2008. *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons*. Cambridge: MIT Press.
- Hutto, D. and Myin, E. 2013. *Radicalizing Enactivism. Basic Minds without Content*. Cambridge: MIT Press.
- Hutto, D. and Myin, E. 2017. *Evolving Enactivism. Basic Minds meet Content*. Cambridge: MIT Press.
- Hutto, D. and Myin, E. 2019. "Deflating deflationism about mental representations." In J. Smortchkova, K. Dolega and T. Schlicht (eds.). *What are Mental Representations?* New York: Oxford University Press, available at https://www.academia.edu/34441867/Deflating_Deflationism_about_Mental_Representation
- Hutto, D. and Satne, G. 2015. "The natural origins of content." *Philosophia* 43 (3): 521–536.
- Jackendoff, R. 1992. *Languages of the Mind*. Cambridge: MIT Press.
- Jacob, P. 1997. *What Minds Can Do*. Cambridge: Cambridge University Press.
- Jacob, P. 2010. "The scope and limits of Chomsky's naturalism." In J. Bricomont and J. Franck (eds.). *Chomsky Notebook*. New York: Columbia University Press: 211–234.
- Lloyd, D. 2003. "Representation, philosophical issues about." In L. Nadel (ed.). *Encyclopedia of Cognitive Science*. Vol. 3. London: Nature Publishing Group: 934–940.
- Milkowski, M. 2015. "The hard problem of content: solved (long ago)." *Studies in Logic, Grammar and Rhetoric* 41 (1): 73–88.
- Morgan, A. and Piccinini, G. 2018. "Towards a cognitive neuroscience of intentionality." *Minds and Machines* 28 (1): 119–139.
- Newen, A., De Bruin, L. and Gallagher, S. 2018. *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.
- Noë, A. 2004. *Perception in Action*. Cambridge: MIT Press.
- Rey, G. 2003. "Chomsky, intentionality, and a CRTT." In L. M. Antony (ed.). *Chomsky and His Critics*. Malden: Blackwell: 105–139.
- Rosenberg, A. 2013. "Disenchanted naturalism." In B. Bashour and H. Muller (ed.). *Contemporary Philosophical Naturalism and its Implications*. New York: Routledge: 13–17.
- Rosenberg, A. 2015. "The genealogy of content or the future of an illusion." *Philosophia* 43 (3): 537–547.
- Roy, J.-M. 2015. "Anti-Cartesianism and anti-Brentanism: the problem of anti-representationalist intentionalism." *Southern Journal of Philosophy* 53 (Spindel Supplement): 90–125.
- Searle, J. 1983. *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT Press.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science. The Case against Belief*. Cambridge: MIT Press.
- Stich, S. 1996. *Deconstructing the Mind*. Oxford: Oxford University Press.
- Thompson, E. 2007. *Mind in Life*. Cambridge: Harvard University Press.

- Titchener, E. 1909. *Lectures on the Experimental Psychology of the Thought-Processes*. New York: MacMillan.
- Varela, F. 1992. "Autopoiesis and a biology of intentionality." In B. McMullin (ed.), *Proceedings of the Workshop "Autopoiesis and Perception"*. Dublin: Dublin City University: 4–14.
- Wheeler, M. 2017. "The revolution will not be optimised: radical enactivism, extended functionalism and the extensive mind." *Topoi* 36 (3): 457–472.

Book Reviews

Clare Chambers, Against Marriage: An Egalitarian Defense of the Marriage-Free State, New York: Oxford University Press, 2017, 226 pp.

In the last few years we have witnessed a remarkable change in social and political attitudes about marriage. After decades of campaigning, the cause for same-sex marriage has scored a string of success. Today, same-sex marriage is legally performed and recognized (nationwide or in some parts) in Argentina, Australia, Belgium, Brazil, Canada, Colombia, Denmark, Finland, France, Germany, Iceland, Ireland, Luxembourg, Malta, Mexico, the Netherlands, New Zealand, Norway, Portugal, South Africa, Spain, Sweden, the United Kingdom, the United States, and Uruguay. Though not legal *per se* in Armenia, Estonia, and Israel, same-sex marriage performed abroad is legally recognized in these countries.

At the same time, the institution of marriage has come under intense scrutiny from political philosophers working within a broad liberal tradition. These philosophers agree with the advocates of same-sex marriage that it is unjust to make marriage available to opposite-sex couples while excluding same-sex couples. But some of these philosophers give only their heavily-qualified support to the same-sex marriage campaigners' fundamental goal—which is to give same-sex couples access to something that closely approximates the current institution of marriage. According to these political philosophers, marriage, in its current form, is incompatible with liberal principles of justice. Thus, on their view, marriage should ideally be either radically reformed or completely abolished. With her *Against Marriage*, Clare Chambers leads the latter group and provides an extremely rich, lucid, and timely argument against state-recognized marriage.

Chambers opens her book with the claim that it is “for everyone, regardless of marital status” (1). Indeed, one of the book's strengths is that the arguments presented can be appealing to all five groups she lists at the outset: “the happily married,” “the happily unmarried,” “the unhappily married,” “the unhappily unmarried,” and “children, whose social wellbeing should not depend on their parents' marital status” (1). But, as Chambers also warns, “this book is *not* for everyone regardless of political conviction” (1). Her argument is driven by a commitment to egalitarianism arising from the feminist belief that “society is deeply gendered, in a way that harms women, and [that] this is wrong” (2). So, social conservatives or religious fundamentalists who wish to retain a particularly sexist and

heterosexist organization of intimacy are likely to be, at best, untouched by the arguments of the book. Nonetheless, anyone with some egalitarian commitments will find plenty of interesting material to dwell on.

The book has two parts, each with three chapters. Part I, entitled *Against Marriage*, defends the negative thesis: state recognition of marriage in any form should be abolished. Canvassing a broad range of philosophical literature on marriage, Chambers argues that state-regulated marriage, both as it currently exists and even when radically transformed, violates core principles of liberalism: those of equality and liberty.

Chapter 1 argues that marriage violates liberalism's commitment to equality by, at best, ignoring and, at worst, perpetuating the sexist and heterosexist foundations of marriage. Marriage is sexist both symbolically and materially. The marriage ceremony wrongs women symbolically when they are expected to wear a white dress to symbolize their virginity and purity, when the father "gives away" his daughter to her husband, and when the minister (or priest) tells the new husband that he can kiss his bride as if she no longer has to consent to such things. Materially, marriage is problematic because of an all-too-common unequal division of labor within marriages, the difficulty for women to remove themselves from abusive marriages, and the continued financial imbalance between men and women which is usually exacerbated in divorce settlements. Moreover, marriage is heterosexist because, in most jurisdictions still, it only permits marriage between one man and one woman. Chambers also argues here that civil unions (or other reformed versions of marriage) "maintain inequality between partnered and unpartnered people" of all sexualities (42).

State-recognized marriage also violates liberty, Chambers argues in Chapter 2, by promoting a particular conception of the good without sufficiently weighty public reasons for doing so. It does so in part because it offers advantages to married couples not available to unmarried couples, such as tax breaks, pension/insurance benefits, visitation rights, and immigration status. Hence, people might choose marriage not because they really want to, but in order to achieve benefits that would not be available to them otherwise. This violates the liberal principle that the state ought to remain neutral between different conceptions of the good life.

In Chapter 3, Chambers considers several liberal justifications for state-recognized marriage including arguments based on the idea (i) that marriage helps to convey social meaning and thus enables better communication, (ii) that marriage can promote gender equality, (iii) that marriage fosters care, (iv) that marriage promotes social stability, and (v) that marriage is necessary for child protection. In each case, Chambers contends, the arguments fail to show that state-recognized marriage is both a necessary and an acceptable means of achieving the public good in question.

Part II of the book, *The Marriage-Free State*, sets out the positive thesis: since in an ideally just state there is no state-recognized marriage, personal relationships still need to be regulated "so as to protect vulnerable parties, including but not only children; so as to regulate disputes over such matters as joint property; so as to establish the rights and duties of third parties; and so as to appropriately direct state benefits and taxes" (115). Personal relationships should be regulated, Chambers holds, in a piecemeal

way based on the practices people engage in. This form of regulation distinguishes Chambers from others who argued for the abolition of marriage.

In Chapter 4, Chambers considers and rejects relationship contracts as a replacement for state-recognized marriage. Though relationship contracts give significant freedom for people to regulate their relationships in a way that best fits their individual needs, they too can undermine both liberty and equality. The parties to the contract, particularly women, may not be free and equal at the moment of making the contract, or they might end up unfree and unequal by agreeing to certain terms, or both. Moreover, relationship contracts can prove to be difficult and even undesirable to enforce: after all, if relationship contracts are to be legally enforceable, the state is invited to intervene on any matter on which the parties have chosen to contract. This would be giving far too much power to the state with respect to our private lives.

In Chapter 5, hence, Chambers makes her own original contribution regarding the question how personal relationships should be regulated in a marriage-free state. She argues that intimate relationships between citizens should be regulated (1) in a *piecemeal* rather than holistic fashion; (2) with a focus on relationship *practices* rather than status; and (3) with the freedom to *opt out* of default regulations rather than to opt in. Such a regulation would mean that there would not be any particular “relationship status,” such as traditional marriage, Tamara Metz’s Intimate Care-Giving Union, or Elizabeth Brake’s Minimal Marriage, that would provide a predetermined set of rights and benefits. Instead, regulations about care, cohabitation, economic and emotional interdependence, and parenting would exist independently from one another. This avoids the assumption that all the goods of being in a relationship are to be achieved in one relationship. The emphasis on relationship practices rather than on relationship status as well as the requirement to opt out instead of to opt in are meant to restrict anyone from passing on any value-judgment and to protect those who, for whatever reason, lack the relevant status.

While arguing against state-recognized marriage, Chambers acknowledges the importance of religious and secular marriage ceremonies and practices to some (perhaps many) citizens, both straight and gay. Rather than critiquing or devaluing the desire to be married (as some marriage critics do), she accepts this desire as legitimate and seeks to accommodate it. Thus, in Chapter 6, Chambers distinguishes the marriage-free state from the marriage-free society and considers the circumstances under which the state might be justified in intervening in private marriages. Here Chambers clearly distinguishes her own position from a libertarian one by focusing on the state’s role in preventing harm to vulnerable parties (such as women and children) and on ensuring that objectionable discriminatory practices are prohibited in the private sector too. Chambers skillfully argues, for example, that religious exemptions from anti-discrimination legislation cannot be upheld by reference to freedom of association. When women are excluded from priesthood and lesbian and gay Catholics are excluded from the rites of marriage, their freedom of association is hindered: they become outsiders, “people who must be excluded for the comfort and bonds of others” (180). However, unlike private clubs, where adults who join them are

aware of the group's commitments, Chambers reminds us, religions are "a lot more like states and other public institutions" (183). They have, in other words, a profound effect on people's lives from birth: some women, gays, and lesbians have been members of their religious communities from childhood. Thus, they are insiders to a community that actively and unjustly harms them by excluding them on the basis of belonging to a certain group (women, gay, and lesbian) for which they had no choice. Similarly, religious law governing marriage (and divorce), such as Jewish religious law that is still upheld by Conservative and Orthodox Jews, must not allow some members to live under a religious authority that harms them without their consent. "The state must do whatever it can," Chambers argues, "to ensure that women and other vulnerable people are not trapped in oppressive religious marriages" (199). Religious freedom, Chambers concludes, cannot include the right of religious leaders to discriminate against members of their *own* religion.

Throughout the book, Chambers displays an impressive familiarity with both feminist and non-feminist scholars of marriage and the family, both critics and advocates of marriage, representatives of a variety of liberal theoretical approaches, as well as with religion and culture. Chambers is also masterful at anticipating and responding carefully to potential objections to her arguments and proposals. Moreover, as a philosopher, Chambers is thoughtful, precise, and meticulous; as a feminist, she is concerned, compassionate, and attentive to the complexity and diversity of people's lives. Nonetheless, there are certain worries about some of her key points.

While there is much to agree with in Chambers' critique of the institution of marriage, there is at least one important concern about her negative argument. The foci of Chambers' argument against state-recognized marriage is, among others, a concern about the unjustified stigmatization (or as she terms it, borrowing from the work of Pierre Bourdieu, "symbolic violence") of single and unmarried people. It is vague, however, what the stigmatization consists in. One possibility is that it involves the negative beliefs some (perhaps most) members of society have about being single or unmarried. If this is so, it is unclear that it is the business of a politically liberal state to eradicate such beliefs. After all, a neutral, liberal state should not be taking sides on disputed questions about what is a good life; that is, individuals should have the autonomy to decide for themselves what is of value. If the state should not be encouraging marriage in any way, it should not be encouraging being single or unmarried either.

Another possibility is that the stigmatization Chambers identifies involves unjustly discriminating against single or unmarried people, or in some other way infringing on their civil and political rights. In this case, a politically liberal state should certainly aim to protect them from such injustices. Yet, it is unclear that the abolition of marriage is necessary or effective to achieve this end. Even without state-recognized marriage, some citizens will still be coupled and others will continue to be single; the latter might well be stigmatized or discriminated against in a marriage-free state.

Although equally informed by her thoughtfulness and attentiveness, Chambers' central proposal—namely, that personal relationships should be regulated through a series of piecemeal regulations governing various rela-

tionship practices—invites some questions too. I am highly sympathetic to her argument against “bundling” different relationship practices: according rights and benefits to all and only those relationships in which economic, affective, domestic, sexual, and reproductive care are intertwined makes a vast array of other relationships invisible, leaving those who may be most in need of support without state protection. Piecemeal regulation of relationship practices may not, however, be as easy as Chambers suggests. Will this approach achieve its desired egalitarian end of securing both liberty and equality without intruding on people’s privacy? Chambers suggests that:

The state does not need to know a great range of details about people’s private lives so as to determine whether they meet the criteria of being in A Relationship. It simply needs to know whether some particular relationship practice applies. In some cases this will be a matter of objective fact: are the people parents? Or, are they both names on the deeds of a property? Or, does one have caring responsibilities for the other? (155)

Liberal states, Chambers rightly notes, typically have procedures in place to determine the answers to these questions; however, these procedures may not always be so determinate or just.

Imagine a now lesbian woman, happily partnered but unmarried, with children from her ex-husband, taking care of children from her new partner, and living in an extended family. Who are the parents to whom in this scenario? Who has caring responsibilities? With the advancement of medical technology even biological parenthood is now in question. Imagine two women seeking IVF treatment at the same hospital. Only one of them becomes pregnant as a result of the IVF treatment, and at some stage during the pregnancy it is revealed that, due to a mistake, the fetus is the other woman’s genetic offspring. The fetus is carried to term by the gestational mother. Or, thanks to mitochondrial replacement therapy and the existence of gestational procreators, a child can have up to four different biological procreators—three genetic and one gestational. Again, who are the parents?

Conflicts in such cases have been settled differently in different jurisdictions; this indicates that when family configurations and methods of reproduction are complex, decision procedures may fall short of providing definitive answers. Even where existing state procedures deliver a *definitive* answer, we cannot assume that those procedures or results are thereby *just*. Indeed, the primary argument of *Against Marriage* is that procedures for allocating rights and responsibilities in state-recognized marriage regimes are fundamentally *unjust* both to many who are married and to many who are not.

Against Marriage is certainly a wonderful addition to a growing philosophical literature on marriage. A review, unfortunately, cannot possibly do justice to all the thought-provoking and rich material found in the book. Hence, I shall conclude by saying that, despite some worries, Chambers’ arguments against state-recognized marriage are powerful, thorough, and timely, as are her critiques of the most influential alternatives found both in philosophy and legal theory. Chambers’ articulate style, philosophical rigor, and respect for her philosophical interlocutors through a fair treatment of their positions make *Against Marriage* an excellent model of how philosophy can be accessible to both professional philosophers and those

with no philosophical training. *Against Marriage* is surely going to compel its readers to think deeply whether the institution of marriage is still viable and about how the regulation of adult personal relationships may have to be altered to meet the demands of justice.

MARKO KONJOVIĆ

Central European University, Budapest, Hungary

Tim Crane, The Meaning of Belief: Religion from an Atheist's Point of View, Cambridge: Harvard University Press, 2017, xiv + 207 pp.

While not intended as an academic treatise on religion, Tim Crane's *The Meaning of Belief* is a thoughtful, interesting reflection on the nature of religious belief in contemporary life. The book is clearly and fluidly written, as one would expect from Crane, and its main theme will be refreshing to those who are weary of the cliché-ridden debates between the New Atheists and their theistic interlocutors. Crane's central thesis is that many of his fellow atheists incorrectly identify religion with a mere set of cosmological and moral propositions, falsely leading them to believe that religious people will tend to change their minds after exposure to the right philosophical or scientific arguments. The book is very rich, and it would be impossible to name all its virtues without resorting to a laundry list; I recommend the title for its expansiveness alone. However, I found Crane's overall argument unconvincing.

Crane asks, "What is religion, and how does it move people?" and accurately responds that a strict, universal definition of religion is probably impossible (2–4). He endorses Durkheim's claim that religions are best understood by following how they developed historically, and then provides his own definition of religion: "Religion, as I am using the word, is a systematic and practical attempt by human beings to find meaning in the world and their place in it, in terms of their relationship to something transcendent" (6). He elaborates upon these points in the next two chapters in terms of "the religious impulse" and the phenomenon of "identification."

One puzzling aspect of this first chapter is Crane's insistence upon the theoretical and practical value of his own definition of religion, despite his admission that religion probably does not have an exhaustive or universal essence. He seems to vacillate between a hard realism, which (apparently) prevents one from calling "socialism, communism, environmentalism, scientism, humanism, secularism, and atheism" religions ("[W]hat would be the point of this?" Crane asks (24)) and a softer anti-essentialism which allows that there is likely no single essence of religion. Crane recognizes that religion is a historically-conditioned category that may not have been used as a concept through much of history; in the end, however, he clearly settles into a realist mode. I wish he had better explained and justified this move. Why is it best to act as if that there is some real phenomenon which lies behind the bundle of characteristics (systematicity, practicality, mean-

ing-seeking, relatedness to the transcendent') which are supposedly concatenated in 'religion', given that these characteristics can be isolated and found in combination with other things (as Crane acknowledges)? I do not believe that Crane successfully captures the phenomenon (if there is one), and this is presuming that the individual items in the 'religious' bundle (e.g. 'transcendence') are well-defined categories themselves, which I do not think Crane proves. I think that the book also overlooks motives which both secularists and religious people might have for wishing to do away with the category of religion at a political level. On the secular front, some might point out that religious organizations are often given special protections and privileges simply because they are 'religious,' and some might see this as unjustified. Similarly, religious people might note that 'religious' arguments are seldom taken seriously in public discourse and might inquire as why these arguments deserve no consideration simply because (if Crane is correct) religion appeals to a normative transcendent order or involves group identity. Crane does not closely address such political considerations, which is unfortunate given their relevance to his overall message.

Other interesting discussions in the chapter concern the nature of the supernatural (Crane does not believe that this category is helpful in distinguishing religion from non-religion, since the category is based on contestable, modern assumptions about nature, and would not distinguish religion from magic (10)); the features of belief ("accessibility to consciousness, connection to action, and the aim toward truth" (16)); and the differences among atheism, agnosticism, and humanism. Crane's points on these topics are never groundbreaking, but they are clear and plausible. Crane's novel thesis—that the New Atheists miss the point of religion—is less plausible; I shall return to this later.

The next chapter discusses an element of religion which apparently differentiates it from other belief-systems: relatedness-to-transcendence or "the religious impulse." Crane cites James, explaining that the religious impulse is the tendency to believe in an unseen order which is the source of normativity (36–37). Crane accepts James' general hypothesis, though notes a difficulty with it: everyone, including the scientist, believes in an unseen order, whether it be an order of gods or quarks. Crane's straightforward admission on this point is admirable, as are his attempts to clarify the difference between the transcendent world of science and the transcendent world of religion. But I do not think that he succeeds, and his thoughts become less and less convincing as the chapter progresses. At points, he seems to disassociate religious transcendence from normativity; at others, he seems to believe that normativity is the distinguishing mark of religious transcendence. Eventually, he attempts to define religious transcendence as involving an inherently unknowable element—this would explain, he speculates, why religious people seem immune to argument (the unknowable cannot be refuted). But what does this make of Locke's substance or Kant's *noumenon*? What if Nagel is right and we can never know what it is like to be a bat?¹ Are substances, *noumena*, and bats religiously transcendent is unknowable in some respect? Are Locke, Kant, and Nagel fundamentally religious thinkers? Anyways, it is not obvious that all religions espouse this

¹ Nagel, T. 1974. "What is it like to be a bat?" *Philosophical Review* 4: 435–50.

sort of transcendence. The ancient Greek and Roman gods, while perhaps awe-inspiring and illusive, could nonetheless be interpreted as immanent parts of the Greek and Roman worlds—and, moreover, Greek and Roman religion was not distinguished from the earthly state by its practitioners. And many in the ancient world, regarded philosophy, not religion, as the path to communion with ineffable Being.

But this is not the sole problem with the Jamesian thesis: belief in a transcendent order which is the source of normativity is not an exclusively or even primarily religious idea. In fact, the normativity of nature, in both its seen and unseen elements, is at the core of Western philosophy—it is of central importance in the work of the Socratics, Cynics, Stoics, and Neoplatonists. It was a core concept throughout medieval philosophy, and it is even accepted by some natural law theorists today. Appeals to nature in ethics, fallacious or otherwise, are common enough that have been labelled an error of thought—noteworthy, ‘the *naturalistic* fallacy’ not ‘the *religious* fallacy.’ I wish the book had addressed these issues more fully, to better provide the reader with a sense of the meaning of religious morality.

That said, Crane makes the largely commonsense claim that for the religious person, harmony with the transcendent order (in the Jamesian sense) consists in living in accord with the divine will—though he does not at all mention the divine intellect, which would have been far more important for many intellectualist mystics, such as Meister Eckhart or Thomas Merton, or for intellectualist thinkers like Aquinas. Crane says that surely the divine will lies at the heart of religion—any person who has thought hard about the nature of belief will see this. Fair enough—I am not aware of many religions which hold God’s will to be *bad*, though it should be noted that some pagan religions treat the gods as capricious forces to be bribed or tamed (perhaps one might label these religions ‘magical’). But, as Crane rightly notes, “The difficulty lies in detailing what this [following God’s will] exactly means” (38). That said, Crane does *not* then expound upon why religious people think God’s will is to be followed, an omission which has the potential for creating serious misunderstandings. For instance, in the absence of further explanation, it might sound as if Crane is suggesting that religion fundamentally involves a commitment to voluntarism or divine-command theory, which it does not. Numerous religious philosophers have non-ambiguously taken on the non-voluntarist horn of the Euthyphro dilemma—or have attempted to dodge the dilemma altogether. Even non-scholarly religious people commonly (in my experience) attempt to justify *why* God thinks various things are good—they do not simply say that God’s commands are good because God commanded them. One need not be a Leibniz scholar to appreciate Voltaire’s scathing critiques (e.g. in *Candide*) of optimistic, naïve religious explanations of how the slings and arrows of everyday life are somehow good and therefore somehow ordained by providence. In short, it is plainly false that it is part of the true essence of religion to simplistically respect God’s will because it is God’s will. But, even supposing that all religious people were voluntarists or divine-command theorists of some sort, there is no reason to suppose that God’s will would be truly ‘normative’ for them in the moral sense. It is perfectly conceivable that one might believe in God and regard God’s commands as to be obeyed with fear and trembling,

because God is the gunman-writ-infinite—if one does not obey, one will be punished. But fear of punishment need not be taken as a revelation of the true, good, or beautiful, as anyone who has seriously reflected on the meaning of these words will know—and many non-scholars, including religious non-scholars, do reflect on them. Overall, the chapter underemphasizes the great diversity of religious thought on normativity.

These objections aside, Crane posits that the ‘religious impulse’ is about seeking meaning in life and that the transcendent order provides meaning for religious people. Crane then asks, ‘What about those who disbelieve in a transcendent order?’ According to the ‘pessimist’ account (which Crane espouses), there is no inherent meaning to life, so we might as well learn to live without such meaning. According to the ‘optimist’ account, on the other hand, one might attempt to explain how, in some way, there can be meaning to life without a transcendent order: one might even argue that the notion of a transcendent order is completely irrelevant to the question of meaning. I found this entire discussion confusing—in fairness, I usually struggle to understand the ‘existential’ (if not semantic) meaning of meaning, which people often (incoherently, in my view) use when they discuss the meaning of life. That said, I must say that this section of the book did little to clarify my confusion. How does God’s will provide existential ‘meaning’ for the religious person? Does God say something is meaningful, and it is meaningful? This is surely too simplistic an account. It does not account for the ‘absurdist’ theist who might find God’s commands arbitrary and meaningless, even if he believes that they must be obeyed for the sake of avoiding Hell. Nor does meaning necessarily relate to morality, it seems, since Crane thinks we can have morality without a meaning to life—though, it is noteworthy that Crane also asserts that, as a ‘pessimistic’ atheist, he would find morality easier to make sense of if there were a transcendent order. I wish Crane would have expanded upon this thought—it would have clarified the vaguer points regarding the relationship between the religious impulse and normativity. Overall, the meaning of ‘meaning’ was unclear to me; but, in fairness, perhaps others will find it more meaningful.

Later in the chapter, Crane states that if religious believers were to take certain problems raised by atheists (e.g. ‘If God created the universe, who created God?’) more seriously, they ought to be worried. Of course, religious believers might retort that if Crane took certain problems raised by theists (e.g. ‘How does one explain fine-tuning?’) more seriously, he ought to be worried. Crane is surely aware of such retorts, and he would not be unduly self-confident if he thought he could convincingly respond to many of them in atheism’s favor. What worries me is that he seems to conclude that religious sentences uttered by ordinary religious believers are not even meant to be *cognitive*—at least, not in the normal sense presumed by science (or common sense), wherein true beliefs are presumed to predict or accurately correspond to some state of affairs in the past, present, or future. According to Crane, when religious people use sentences such as ‘God created the world,’ they are (usually?) using them to express something other than a scientific belief, because the sentences would obviously be false if interpreted scientifically. Crane compares religious sentences to historical ones—historical claims are true or false, but do not involve simple laws or

predictions, like science's sentences do. But Crane simultaneously uses a historical example to point out that religious people are not concerned with 'scientific' prediction; Jesus apparently predicted his own return, which did not happen, and Crane concludes that religious people must not be worried about prediction, since they are not bothered by Jesus' failed prophecy. But in my experience, religious believers are bothered by it; some who become aware of the difficulty go into denial, others become atheists, and yet others choose to adopt the traditional Preterist reading of the Bible, wherein Jesus' prophecies *were* predictive, because they correctly predicted the destruction of the Second Temple in 70 A.D. One way or another, it seems unnecessary to posit a uniquely 'religious' mode of semantics in order to account for the data, and the meaning of Crane's non-scientific 'meaning' of religious sentences remains vague.

That said, I suspect that many religious people would be confused by Crane's thesis regarding the meaning of their belief; that is, I doubt Crane's thesis holds on a statistical level. The average religious person may not have arguments for God's existence or may not understand that his or her arguments for God's existence are fallacious, and he or she may not be interested (on an average day) in responding to critique, and sometimes his or her reticence to consider atheist arguments might be unjustified—but the same sort of things go for most people regarding most topics. We all make mistakes and we all dislike having these pointed out, because it hurts our reputations when we are exposed as wrong, and because we like to think that we are reasonably correct about the world and can have confidence that our beliefs are accurate. It is very threatening to be told that one's core beliefs are misguided, both for the common religious person and for the sophisticated philosopher. (After all, who genuinely enjoys the peer-review process, even if he or she is genuinely committed to the truth?) This does not make the essence of religion something vaguely emotive like honking for Jesus or cheering 'Hoorah for the mysteries of faith!' To be sure, *some* religious people regard expressions like 'Jesus rose from the dead' as non-factual (in the ordinary, historical, or 'scientific' sense) but as (somehow) still meaningful. But many religious people—probably most—do not. In my own experience, average religious people tend to believe that the propositions they express in their religious sentences are true in the 'scientific' sense, and many of them would be greatly bothered by the suggestions that these propositions are not factually correct or that they are not factually meant. Ironically, the religious 'non-cognitivist' is more likely to be the sort of sophisticated *theologian* whom Crane does not wish to discuss in this book. If religion is the opiate of the masses, the obscure search for religious 'meaning' in the face of the facts is probably a luxury of the bourgeois. But this is a matter of statistical and sociological hypothesis.

Other interesting discussions in the second chapter concern the cognitive science of religion (Crane thinks that contemporary accounts are too simplistic), the relationship between science and religion (I find Crane's account difficult to accept, as may be inferred from the above) and the meaning of faith (it is not dogmatic certainty, but something like a commitment to understanding the transcendent mystery of reality in the face of uncertainty).

In the third chapter, Crane turns to the phenomenon of ‘identification,’ which is the second pillar of religion. Crane correctly points out that religions are groups of people who function according to the laws of psychology and, and are not mere collections. But, as in the section on religious semantics, Crane draws distinctions which I do not believe exist. For instance, Crane suggests that aspects of religion like fasting and prayer are not explainable as the results of religious cosmological or moral belief. For, surely, if they were expressions of either of these, they would be expressions of the latter—but morality clearly consists in how one treats others. Therefore, fasting and prayer must be part of a third, sui-generis category he calls ‘religious practice.’

This argument is dubious. Recent work in moral psychology² suggests that practices like fasting or prayer may very well be expressions of intrinsic, evolutionarily conditioned components of human morality—for instance, they may express innate moral tastes for ‘purity’ or ‘respect,’ even if these would not be recognized as truly ‘moral’ concerns by Western liberals. And simple observation reveals that many religions have prohibitions and taboos which have no clear relationship to ‘how we treat others’ in the Western liberal sense, but which are still clearly recognized as ‘moral’ injunctions by believers. For instance, the Catholic Church treats non-procreative sex as gravely immoral—not because non-procreative sex necessarily leads to physical, emotional, or societal harm, nor because it necessarily involves ‘using’ another person or violating another’s autonomy *per se*, but rather because the Church holds that non-procreative sex is ‘unnatural,’ in the Aristotelian sense. In fairness, it has been argued that Christianity borrowed this particular taboo from Stoicism, not the Bible or Church tradition³—but, borrowed or not, the taboo is still, in some sense, a matter of belief for this religious group, and, doubly contrary to Crane’s thesis, it is both a moral *and* cosmologically-based belief, since it only has any plausibility given that the universe is understood to function in a certain (i.e. teleological) way. Many other examples could be given as to how to reduce ‘religious practice’ to moral and cosmological belief—in fact, it is not clear that in-group identification is not regarded as a moral affair by religious communities. People who are not Western liberals often regard identification (or ‘loyalty’ and ‘tradition’) as a moral concern which pertains to the overall functioning of the universe (consider the Indian caste system and its relation to Dharma). It is worth remembering that many ancient Christians were persecuted by Romans for treason (and ‘atheism’) because they refused to participate in *patriotic* and religious rituals, such as paying homage to Caesar or the gods, because patriotism, religion, and loyalty were regarded as the same thing by many ancient peoples. (Consider that Imperial Rome’s ‘official’ founding figure was ‘Pius Aeneas.’) To this day, many non-religious citizens (including atheists) find it distasteful to burn national flags or to refuse to rise during national anthems—whether Western liberals acknowledge loyalty or tradition to be a

² See, for instance, Graham, J., Haidt, J., and Nosek, B. A. 2009. “Liberals and Conservatives Rely on Different Sets of Moral Foundations.” *Journal of Personality and Social Psychology* 96 (5): 1029–1046.

³ See Noonan Jr, J. T., and Noonan, J. T. 2012. *Contraception: A History of Its treatment by the Catholic Theologians and Canonists*. Cambridge: Harvard University Press.

truly 'moral' affair or not. In sum, Crane's ground for positing the category of *sui generis* 'religious practice' is weak—religious taboos and traditions can probably be accounted for in moral and cosmological terms.

Crane also discusses the meanings of patriotism, pride, and shame, and briefly discusses identification in the context of Rawls. He argues that the sort of religious upbringing which the New Atheists call child indoctrination is more akin to the sort of cultural adjustment one receives in the family, and he is surely correct on this. Further on, he explains that the sorts of meaningful practices, traditions, and histories which lend meaning to religion also lend meaning to secular pursuits like academia—and again, he is surely correct. Crane completes the chapter with a discussion of Durkheim's notion of the Sacred, with Sacred things symbolically uniting the religious impulse with the phenomenon of identification through their intentionality. Skeptics of Durkheim will be reticent, but it is hard to see how Crane's basic point could be inaccurate.

The fourth chapter covers the topic of religious violence, and Crane's main points are very plausible. Crane notes that many supposedly religious conflicts have cultural, ethnic, and political undertones such that it is difficult to account for them simply in terms of doctrinal differences; that is, while differences in cosmological or moral beliefs can make a difference, the element of identification and the numerous other factors which generally fuel human conflict are probably more important. Crane understates the roles which particular beliefs do have in causing conflict (the *filioque* clause really did lead to schism), but the main point of the chapter (that religious conflicts are human conflicts) is well taken; Crane has done a great service in highlighting it.

If the first chapter began with a large question ('What is religion?'), the last chapter begins with an even larger assertion: that religion is one of "the main drivers behind world events are religion and nationhood." That said, one must learn to live with religion, because it is "wishful thinking" to suppose that religion will go away (163). To my mind, the difficulty with these assertions is that the overall secularization of many European countries seems to clearly contradict it. Though large percentages of the population do remain nominally religious in even the most atheistic European countries (e.g. France and the Czech Republic), there is little denying that religion (as defined by Crane) has become less important in everyday life. With a fair amount of data available regarding these matters I wish Crane could have argued for his claims more empirically.

The remainder of the final chapter discusses the meaning of tolerance. Being tolerant is not relativistic or paternalistic, nor does it imply respect for the beliefs one merely tolerates. By definition, toleration implies disapproval—one could not tolerate religious beliefs if one liked them or thought that they were true. The most interesting, though also most underdeveloped, section of this chapter is about the limits of tolerance. At a political level, Crane believes that we should not tolerate religious behavior beyond the rule of law, though what constitutes a reasonable boundary between law and religious liberty remains vague. This, to me, would have been a fascinating section to have more fully fleshed out. Crane spends most of the remainder of the section addressing how, on a personal level, an atheist ought

to treat legal religious behavior. ‘It depends on the circumstances,’ seems to be the answer. Sometimes, it may be morally commendable to even participate in religious rituals, out of respect for one’s religious neighbors—for instance, by agreeing to wear a skull-cap at a Jewish friend’s funeral. In other cases, one might rightly protest offensive (but legal) religious behaviors, such as the prohibition of women from the priesthood in some churches. But one should not argue against objectionable practices with philosophical or scientific arguments—these are unlikely to have an impact, since religious folk are not generally receptive to truth, even if many are reasonable and highly educated. Tolerant dialogue, which involves understanding what religion is about (an attempt to find meaning through identification with some group which talks about the transcendent), is preferable. That said, I did not find a recommendation as to how dialogue might persuade religious people to avoid offensive or immoral practices, given that philosophical or scientific arguments are off the table.

In summary, the difficulties of Crane’s book lie not in its content, but in the ways in which it could have been more fully expanded in order to uphold his main criticism of the New Atheists. Crane has made fascinating psychological and sociological claims regarding religion—that religion will not go away, that religious people are unreceptive to rational arguments, that this is because of group identification and belief in an unknowable transcendent order. But Crane does not convincingly argue for *why* it is solely religious people who have their access to reason blocked through peculiar beliefs or in-group identification. As Crane acknowledges, there are many forms of belief and many forms of in-group identification—why is religion so special? Why have so many rational, non-religious (in Crane’s sense of the term) thinkers in the philosophical tradition—Plato, Aristotle, the Stoics, Plotinus—held beliefs in a transcendent, normative order, if such beliefs are inherently ‘religious’? Is Aristotle’s ‘theology’ just so much non-cognitive expression of in-group identification? On a more scientific note, Crane does not always back up his suggestions up with empirical data, which is problematic considering that common-sense and historical intuition contradict them. For instance, Crane largely ignores the fact that religious people *do* occasionally (I do not know how often—that is a statistical question) change their minds in the course of argument—this is often how atheists are made (and vice versa). Nor does Crane pay attention to the fact that some religious people genuinely *do* regard religion as a series of cosmological and ethical beliefs, such as the theologians whom Crane straightforwardly admits he will not discuss in the book. Overall, Crane’s view of ordinary religious folk comes across as insufficiently empirical and rather pessimistic—while the New Atheists at least credit religious people with the ability to change their false beliefs, Crane denies that the average religious person is interested in truth. Whether accurate or not, the claim is empirical and requires empirical support. These issues aside, Crane’s book provides food for thought, and creates plenty of ground for future research.⁴

GREGORY FRISBY

Central European University, Budapest, Hungary

⁴ I thank Mark Frisby and Will Zimmerman for comments on drafts of this review.

Croatian Journal of Philosophy is published three times a year. It publishes original scientific papers in the field of philosophy.

Croatian Journal of Philosophy is indexed in *The Philosopher's Index*, *PhilPapers*, *Scopus*, *ERIH PLUS* and in *Arts & Humanities Citation Index (Web of Science)*.

Payment may be made by bank transfer

SWIFT PBZGHR2X

IBAN HR4723400091100096268

Croatian Journal of Philosophy is published with the support of the Ministry of Science and Education of the Republic of Croatia.

Instructions for Contributors

All submissions should be sent to the e-mail: cjp@ifzg.hr. Submissions must be in English and formatted to be double-spaced with suitably wide margins, an A4 page size, and automatic page numbering.

Articles are normally no longer than 8,000 words of main text including bibliography. The Journal will consider longer papers, but, once these limits are exceeded, authors should bear in mind the editorial policy that the acceptance bar raises with increasing length.

Manuscripts should be compiled in the following order: cover page; title; abstract (not exceeding 200 words); keywords (3 to 6); main text; appendices (as appropriate); references.

All the authors of a paper should include their full names, affiliations, postal addresses, telephone and fax numbers and email addresses on the cover page of the manuscript. If a paper is co-written, one author should be identified as the Corresponding Author. The cover page must be submitted as a separate document. All submitted manuscripts must be prepared for blind review, with revealing acknowledgements and self-identifying references removed.

Sources are cited in the text by the author's last name, the publication date of the work cited, and a page number if needed, e.g. (Barber 2007: 324). Full details appear in the reference list in which the year of publication appears immediately after the author's name:

Barber, A. 2007. "Linguistic Structure and the Brain." *Croatian Journal of Philosophy* 21 (7): 317–341.

Williamson, T. 2013. *Identity and Discrimination*. Oxford: Wiley-Blackwell.

The publication of a manuscript in the *Croatian Journal of Philosophy* is expected to follow standards of ethical behavior for all parties involved in the publishing process: authors, editors, and reviewers. The journal follows the principles of the Committee on Publication Ethics (<https://publicationethics.org/resources/flowcharts>).

ISSN 1333-1108



9 771333 110001